Robust Volatility Estimation for Multiscale Diffusions with Zero Quadratic Variation

by

Theodoros Manikas

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Statistics

April 2018

THE UNIVERSITY OF WARWICK

Contents

List of Tables

List of Figures

Acknowledgments

First and foremost I would like to thank my supervisor, Dr. Anastasia Papavasiliou, for all her assistance and guidance. I will always be grateful to her for the opportunity she gave me to work in an ideal and stimulating environment. I am also grateful to the Warwick Statistics department for funding my PhD studies.

Special thanks goes to my friends and flatmates Pantelis and Panayiota who made my life easier and my free time much more enjoyable. They will be one of the reasons to remember Warwick as a pleasant experience in my life.

Also, I would like to express my deepest gratitude to my family members, my parents, Yiannis and Dimitra, my sister Antigoni, and my aunt, Lola. I never would have made it here without their unconditional love, support and encouragement.

Last but not least, I owe my lovely thanks to one more member of my family, my partner Maria. She was my main source of encouragement and the reason that I kept on track in difficult times. Throughout these years she was always by my side even when I was not the most enjoyable person.

Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself except when stated otherwise and has not been submitted in any previous application for any degree.
Theodoros Manikas, April 2018.

Abstract

This thesis is concerned with the problem of volatility estimation in the context of multiscale diffusions. In particular, we consider data that exhibit two widely separated time scales. Fast/slow systems of SDEs that adopt a homogenized SDE are employed to model such data. The problem that one is confronted with, is the mismatch between the multiscale data and the homogenized SDE. In this context, we examine whether if by using the multiscale data, the diffusion coefficient of the homogenized SDE can be estimated. Our proposed estimator consists on subsampling the initial data by considering only the local extremals to overcome the issue associated with the underlying model. We provide both theoretical and numerical heuristics, suggesting that our proposed estimator when it is applied to multiscale data of bounded variation is asymptotically unbiased for the volatility coefficient of the homogenized SDE. Furthermore, for the particular example of a multiscale Ornstein–Uhlenbeck process, the numerical results indicate that the L_2–error of our estimator is very small. Moreover, we illustrate situations where the proposed estimator can also be used for multiscale data with bounded non–zero quadratic variation.

CHAPTER 1

Introduction

In many application areas, it is often the case that the most accurate models to describe the dynamics of a physical phenomenon are multiscale in nature. Such situations can be met in the field of molecular dynamics (Schlick, 2010), in atmosphere/ocean science; see, for example, Majda et al. (2001, 2006) and Katsoulakis et al. (2004, 2005) and in network traffic data (Abry et al., 2002). Also, especially in the fields of econometrics, high–frequently observed financial data exhibit multiscale characteristics in the sense that different features are associated with different time scales. These features are usually described by the term market microstructure noise, which contains all different types of market inconsistencies such as non-synchronous trading and bid–ask spread. Tsay (2005) described each of these effects and give a comprehensive review.

Finding a coarse–grained model that can effectively describe the dynamics of the initial multiscale model is a very popular problem among applied mathematicians. This is mainly due to the fact that such models are much more efficient to use in practice. Once the coarse–grained model has been extracted, the corresponding free parameters are needed to be estimated by fitting the model to the data. In this framework, the problem that one is confronted with is the mismatch between the coarse–grained model and the data generated by the full multiscale system.

The natural way to describe data with multiscale character is to employ multiscale diffusions. There are two types of multiscale diffusions that are usually employed, both considered as fast/slow systems of stochastic differential equations (SDEs), the averaging for SDEs and the homogenization for SDEs. Here, we introduce these two types similarly to Pavliotis and Stuart (2008). For the averaging one has

$$dx = f_1(x,y)dt + a_0(x,y)dU + a_1(x,y)dV, \qquad x(0) = x_0, \tag{1.1a}$$

$$dy = \frac{1}{\epsilon}g_0(x,y)dt + \frac{1}{\sqrt{\epsilon}}\beta(x,y)dV, \qquad y(0) = y_0, \tag{1.1b}$$

and for the homogenization,

$$dx = \left(\frac{1}{\epsilon}f_0(x,y) + f_1(x,y)\right)dt + a_0(x,y)dU + a_1(x,y)dV, \quad x(0) = x_0, \tag{1.2a}$$

$$dy = \left(\frac{1}{\epsilon^2}g_0(x,y) + \frac{1}{\epsilon}g_1(x,y)\right)dt + \frac{1}{\epsilon}\beta(x,y)dV. \qquad y(0) = y_0, \tag{1.2b}$$

where $0 < \epsilon << 1$ denotes a small parameter controlling the scale separation, and U, V are independent standard Brownian motions. In both cases, the process x should be viewed as a process evolving on a timescale which is much slower that the timescale of y. For this reason, the process x will be usually referred as the slow variable of the system whereas y as the fast one. As mentioned earlier, the objective is to obtain a coarse–grained model describing the dynamics of the slow variable x. In Pavliotis and Stuart (2008), the reader can find a comprehensive study of how using averaging or homogenization techniques it is possible to show that the slow process converges weakly (in the limit of $\epsilon \to 0$) to a process X solving an appropriate SDE

$$dX = F(X)dt + \Sigma(X)dW, \qquad X(0) = x_0. \tag{1.3}$$

A brief review of the main concepts of this theory for the homogenization case will be presented in Chapter 2. Although this theory allows us to approximate the slow dynamics of the full multiscale system by a smaller dimension process, it is often the case that the functional form of the full multiscale system is unknown. Consequently, the dynamics of the corresponding limiting diffusion process are also unknown. Naturally, the question that rises is how to estimate the unknown parameters of the limiting diffusion process by a data–driven strategy when the available data are observations from the full multiscale system. Apart from the standard problems in parameter estimation problem for SDEs (see Iacus (2009)) in one is also confronted with the mismatch between the multiscale observations and the single–scale limiting diffusion process for which we want to perform parameter estimation.

1.1 Parameter Estimation for Multiscale Diffusions

The parameter estimation problem in the context of multiscale diffusions has been examined extensively by several authors, see for example Pavliotis and Stuart (2007); Sykulski et al. (2008); Papavasiliou et al. (2009); Papavasiliou (2011); Krumscheid et al. (2013); Krumscheid (2014a); Papanicolaou and Spiliopoulos (2014); Kalliadasis et al. (2015); Krumscheid et al. (2015); Gailus and Spiliopoulos (2017a,b); Papanicolaou and Spiliopoulos (2017). A review of these papers will follow in the following sections.

The statistical estimation problem of our interest is to use data–driven techniques to fit data from Eq.(1.1) or Eq.(1.2) to an SDE of the form

$$dX = F(X;\theta)dt + \Sigma(X)dW, \qquad X(0) = x(0), \tag{1.4}$$

where θ is an unknown parameter. The standard approach for the estimation of the drift and diffusion coefficient in an SDE of the form (1.4) is the maximum likelihood estimator for the drift and the quadratic variation for the diffusion coefficient (see Bishwal (2008); Kutoyants (2013); Liptser and Shiryaev (2013)). However, it turns out that due to the small effects in the observations these estimators are not helpful. in the context of multiscale diffusions. In fact, due to the small effects in the observations these estimator are biased (see Theorem 1.1 in Pavliotis and Stuart (2007)). Also, Krumscheid (2014b) uses a simple example to illustrate the failure of these methods.

In the following subsection we are going to give an overview of these estimators.

1.1.1 MLE and Quadratic Variation

Following the work in Papavasiliou et al. (2009), we assume that Σ in Eq.(1.4) is uniformly positive–definite on $\mathcal{X}$, Eq.(1.4) is ergodic with invariant measure $\nu(dx) = \pi(x)dx$ at $\theta = \theta_0$ and that

$$A_\infty := \int_{\mathcal{X}} \left(\Sigma(x)^{-1}F(x) \otimes \Sigma(x)^{-1}F(x)\right) \pi(x)dx \tag{1.5}$$

is invertible. Then, given data $\{z(t)\}_{t\in[0,T]}$ an application of the Girsanov theorem suggests that the log likelihood for θ satisfying Eq.(1.4) is given by

$$\mathbb{L}(\theta; z) = \int_0^T \langle F(z;\theta), dz\rangle_{\alpha(z)} - \frac{1}{2}\int_0^T |F(z;\theta)|^2_{\alpha(z)}\, dt, \tag{1.6}$$

where

$$\langle p, q\rangle_{\alpha(z)} = \left\langle \Sigma(z)^{-1}p, \Sigma(z)^{-1}q\right\rangle. \tag{1.7}$$

Then, the maximum likelihood estimator (MLE) of θ given z is

$$\hat{\theta}(z) = \mathrm{argmax}_{\theta}\mathbb{L}(\theta; z). \tag{1.8}$$

The quadratic variation on the other hand, for data given on finely spaced partition, $\pi_{(n)}([0,T])$, is

$$\hat{\Sigma}(z) = \lim_{\mathrm{mesh}\pi_{(n)}([0,T])\to 0} \sum_{i=0}^{n-1} \left(z(t_{i+1}) - z(t_i)\right)^2. \tag{1.9}$$

In Pavliotis and Stuart (2007), for a Brownian motion in a two–scale potential example, theoretical (Theorem 1.1) and analytical results illustrated the fact that both the MLE for the drift and the quadratic variation (QV) for the diffusion led to biased estimations. It was shown though, that subsampling the available data at an appropriate rate between the two characteristic time scales of the full system can lead to the accurate estimation of both the drift and the diffusion coefficient of the corresponding coarse–grained model. In Papavasiliou et al. (2009), the problem of drift estimation was considered but in a more general framework. It was shown that the MLE can be effectively used in the averaging case without subsampling but for the homogenization one should subsample the available data to reduce the bias. An overview of these papers can be found in Pavliotis et al. (2008). The techniques developed there were applied to the problem of estimating eddy diffusivities from noisy Lagrangian observations in Cotter et al. (2009). Related work on parametric inference for multiscale data based on subsampling the data can be found in Sykulski et al. (2008); Crommelin (2012). Finally, Spiliopoulos and Chronopoulou (2013), Gailus and Spiliopoulos (2017a,b); Papanicolaou and Spiliopoulos (2017) focused on the MLE for multiscale problems in the case of vanishing noise intensity. The main idea starting from Imkeller et al. (2013) and Papanicolaou and Spiliopoulos (2014) was to apply nonlinear filtering to the initial multiscale model to obtain a reduced dimension model and prove that the MLE corresponding to the latter produces consistent estimators.

Although in theory, subsampling the available data may lead to consistent estimations for the parameters of the limiting equations, the question of how to find an optimal subsampling rate remains. Studies concerning the determination of this rate for Gaussian processes can be found in Azencott et al. (2010, 2011). In cases where the model describing the initial multiscale data is unknown, it is infeasible to find this optimal rate and it is also consists the MLE approach inappropriate.

In the context of the diffusion estimation problem for a simple Ornstein–Uhlenbeck multiscale system, Papavasiliou (2011) tried to overcome this problem by introducing the p–variation estimate, a notion relying in the theory of rough paths Lyons (1998); Lyons and Qian (2002). It was shown that the proposed estimator which provided a non–parametric way of subsampling the data was asymptotically unbiased and furthermore its L_2–error performance was better than the one proposed in Pavliotis and Stuart (2007); Papavasiliou et al. (2009). Apart from the fact that this methodology was developed in the context of a particular example it also lacks of practical usability. The result was proven in a pure theoretical and continuous time framework and it is not easily interpretable in a real–life framework. For example, given data how could one compute their p–variation, i.e., the supremum over all possible partitions? The aim of the work presented here is to address this issue and to extend this methodology to a more general framework.

Other methodologies in the direction of addressing the issue of the optimal subsampling rate have been presented in Krumscheid et al. (2013); Krumscheid (2014a); Kalliadasis et al. (2015). Based on the computational studies in Krumscheid et al. (2013); Krumscheid (2014a) and Kalliadasis et al. (2015) developed a methodology on obtaining a functional relation between the unknown parameter θ and the statistical properties of the model in Eq.(1.4). Then, an estimator of θ was derived via the best approximation of a system of equations. The resulting estimator was shown to be robust with respect to weak pertubations.

The problem of parameter estimation and in particular the diffusion estimation problem for data possessing at least two widely separated characteristic time scales has also been very active in the field of the econometrics and in particular when studying high frequency data. In the following section we present the relevant literature in this field and we highlight the modeling differences but also the similarities in the estimation approach.

1.2 Volatility Estimation for High–Frequency Financial Data

High–frequency financial data are defined as intra–day data observed on the prices of financial assets. Nowadays, the advances in technology allow the wide use of high–frequency financial data from different financial markets and from individuals to buy and sell. This lead to an increasing demand for better modelling and statistical inference regarding the price and volatility dynamics of the assets.

In financial applications, the most common modelling approach is the diffusion process and the model is described as follows:

Let $X(t)$ be the price process of a security (e.g. a stock price), then it is assumed that the log process $Y(t) := \log X(t)$ follows an Itô process,

$$dY(t) = \mu(t, Y(t))dt + \sigma(t, Y(t))dW(t), \qquad Y(0) = Y_0, \tag{1.10}$$

where W is a standard Brownian motion and the functions $\mu(t, Y(t)) : [0, T] \to \mathbb{R}$ and $\sigma(t, Y(t)) : [0, T] \to \mathbb{R}$ are called the drift and the diffusion coefficient (or volatility) respectively. Under local Lipshitz and linear growth conditions on the functions μ and σ, a unique strong solution for the Eq.(1.10) can be derived. This solution is called diffusion process and it is well known that it is a semimartingale. Semimartingales is a natural class of processes for modelling securities in a context that does not allow arbitrage opportunities, see Karatzas and Shreve (2012).

The statistical estimation problem of our interest is the estimation of the diffusion coefficient (volatility). Various methodologies have been developed in the past several years to estimate the integrated volatility. For an overview see the survey by Andersen et al. (2002). The classical non–parametric approach to estimate the integrated volatility is via the quadratic variation of the observations from model (1.10). The latter is usually referred as "realised variance" or "realised volatility". In fact, classical stochastic processes theory suggests that as the sampling frequency increases, the estimation error of the realised variance should be diminished (Karatzas and Shreve, 2012). Based on this, naturally one would believe that in a high–frequency context the realised variance would perform with a negligible estimation error.

However, empirical studies suggest that the "realised variance" is biased for high–frequency data. There are real–data studies in literature indicating that the realised variance does not converge as the sampling frequency increases. In fact it seems to go to infinity (Brown, 1990; Hansen and Lunde, 2006) and more general, for the appropriateness of the realised variance see Barndorff-Nielsen and Shephard (2002); Andersen et al. (2003) In particular, the bias problem is apparent from volatility signature plots, namely plots of realised volatility versus alternative sampling frequencies, that were introduced by Andersen et al. (2000). The reason for this "anomaly" is the fact that the realised variance yields a perfect estimate for the volatility in the hypothetical scenario where prices are observed in continuous time and without measurement errors (Jacod and Shiryaev, 2013). In practice, this is not the case and Awartani et al. (2004) described the fact that there is noise contaminating the data. Also, a very recent study on testing whether a high–frequency data sample can be treated as reasonably free of market microstructure noise at a given sampling frequency has been presented in Ait-Sahalia and Xiu (2017). Based on the above heuristics we may conclude that Eq.(1.10) is not the appropriate modeling approach for high–frequency data and one should seek for alternative models.

The volatility estimation problem for financial data that are contaminated by the market microstructure noise has been studied extensively in literature. Indicatively see Zhou (1996); Corsi et al. (2001); Aït-Sahalia et al. (2005); Zhang et al. (2005); Barndorff-Nielsen et al. (2008); Bandi and Russell (2008); Jacod et al. (2009) and references therein. The common modelling approach in these studies is the following model

$$Y(t) = X(t) + \epsilon(t), \tag{1.11}$$

where X follows Eq.(1.10) and ϵ stands for the noise around Y.

There are five main approaches in literature for the estimation of the integrated volatility under the presence of microstructure noise:

- autocovariance based introduced by Zhou (1996)
- subsampling introduced by Zhang et al. (2005)
- realized kernel methods introduced by Barndorff-Nielsen et al. (2008a,b)
- likelihood approach introduced by Aït-Sahalia et al. (2005)
- pre–averaging techniques introduced by Podolskij et al. (2009); Podolskij and Vetter (2009).

A very first approach for addressing the estimation problem was introduced by Zhou (1996) based on autocovariance under the assumption of constant volatility. Then a decade later, Hansen and Lunde (2006) extended that work to stochastic volatility.

However, the main disadvantage of this approach is the inconsistency.

A non-parametric approach based on subsampling was proposed by Zhang et al. (2005). These authors proposed the two–scale realised volatility estimator which was the first consistent estimator of the integrated volatility under the presence of microstructure noise. An extension of this estimator can be found in Zhang et al. (2006) with the multiscale volatility estimator and later on Aït-Sahalia et al. (2011) improved the efficiency of the estimator. In fact, it was shown that the multi–scale estimator converges to the true value of the integrated volatility at the rate of $n^{1/4}$ whereas the two–scale converges at a rate of $n^{1/6}$.

The realized kernel methods developed in Barndorff-Nielsen et al. (2008a) and their proposed estimator was also consistent and achieved a convergent rate of order $n^{1/4}$ as well. Same authors, Barndorff-Nielsen et al. (2008b), extended their work for non-negative form of realised kernels where now the convergent rate was of order $n^{1/5}$ but had the advantage of a non-negative kernel with probability one, which is generally not the case for the other estimators available in the literature. They derived the limit distribution under various assumptions on kernel weights and Barndorff-Nielsen et al. (2009) applied the methodology of realised kernels in practice where they identified some challenges based on the features of high frequency data and focused on the crucial problem of the bandwidth selection for kernel methods. Recently, Barndorff-Nielsen et al. (2011) combined the two very popular non-parametric approaches of subsampling and realised kernels and concluded that the combination of the two is highly advantageous for estimators based on discontinuous kernels, had no effect on kinked kernels and a negative effect for smooth kernels as it had as a result the increase of asymptotic variance.

The likelihood approach has been initiated by Aït-Sahalia et al. (2005) assuming constant volatility in order to be able to perform maximum likelihood estimator. Xiu (2010) extended this approach to allow stochastic volatility. The latter author demonstrated that the parametric likelihood approach is consistent, efficient and robust with respect to stochastic volatility. Also it has the same model-free feature as the non-parametric approaches i.e. subsampling and realised kernels with the main advantage of the finite sample performance.

Finally, the pre–averaging methods have been developed by Podolskij et al. (2009); Podolskij and Vetter (2009) and are very popular methods for mitigating microstructure in high frequency data. Hautsch and Podolskij (2013) and Christensen et al. (2014) based their work on the pre–averaging methods and took into account jumps.

Jumps are an extension of the model and according to discussions in Jacod and Protter (2012) and Aït-Sahalia and Jacod (2014) both noise and jumps can have an impact on the statistical significance. Recently Mykland and Zhang (2016) extended the pre–averaging method and combined it with M–estimators. This methodology is robust against the noise and the jumps while averaging the continuous part.

1.3 Aim and Objectives

The aim of this thesis is to develop methodology for estimating the diffusion coefficient of a homogenized SDE when the data are modelled via multiscale diffusions.

We will assume data from a multiscale system of the form given in Eq.(1.2). Under the appropriate assumptions, model (1.2) converges in law to a limiting diffusion process whose drift and diffusion coefficients can be found in Pavliotis and Stuart (2008, Chapter 11) or in Bensoussan et al. (2011). The work in this thesis differs from previous research in the area because it does not require the explicit value of the scale separation parameter to construct the proposed estimator.

Our aim is to demonstrate theoretical and numerical results to show the efficiency of our proposed estimator for the diffusion (volatility) coefficient of the corresponding limiting diffusion process.

1.4 Outline of the Thesis

The remainder of the thesis is organised as follows.

In Chapter 2, we review the tools for obtaining coarse–grained model of lower dimension from multiscale diffusions. Chapter 3 introduces our proposed estimator and how it is computed in practice. In Chapter 4, we examine the performance of our estimator on data generated by a simple multiscale system with zero quadratic variation whose fast dynamics are described by a Ornstein–Uhlenbeck process. We provide a solid theoretical proof that this estimator is asymptotically unbiased for the diffusion coefficient of the corresponding homogenized SDE. Our theoretical results are further supported by numerical experiments which also illustrate that its performance is very satisfactory. Chapter 5 presents the extension of our work to more general models and we again prove theoretically that under the appropriate assumptions our estimator is asymptotically unbiased for multiscale models with zero quadratic variation. In Chapter 6, we present numerical studies on different models to illustrate our methodology. In Chapter 7, we examine multiscale models of bounded non–zero quadratic variation. Initially, we compute the value of the proposed estimator when it is applied to Brownian motion paths and we illustrate numerically that our estimator tends to this quantity in limit and not to the desired homogenized coefficient. However, we identify situations where the proposed estimator can be used to obtain correct estimates. Finally, in Chapter 8, we summarise the research contributions in this thesis, discuss their implications and provide future research directions.

CHAPTER 2

Obtaining Coarse–Grained Models from Multiscale Diffusions

In this chapter, the main concepts and methods used in this thesis are introduced. In particular, we review the methodology of extracting a coarse–grained models from multiscale diffusions. We describe two classes of multiscale diffusion processes, the averaging and homogenization for SDEs. However, the main focus of this thesis is the homogenization case as it the most popular in the literature for the parameter estimation problem. The content of this chapter relies on Pavliotis and Stuart (2008).

2.1 Introduction

In this thesis we examine data exhibiting a multiscale character and in order to model such data we employ a diffusion process. As an example, consider data with different characteristic time scales, i.e. phenomena that are characterised by the presence of processes occurring along different time scales. We assume that these different time scale are controlled by a scale separation parameter $\epsilon << 1$ and by using standard techniques, under the appropriate assumptions, we can derive as $\epsilon \to 0$ a coarse–grained diffusion model. This model can effectively describe the dynamics of the original multiscale model. The particular statistical problem of our interest is to estimate the diffusion coefficient of the limiting diffusion process.

2.2 Averaging and Homogenization for SDEs

To generate data with a multiscale character, the following two classes of multiscale diffusion processes have been employed in the papers reviewed in Chapter 1. The first is averaging for SDEs, where the data are generated for a system of SDEs with the following form

$$dx(t) = f_1(x(t), y(t))dt + \alpha_0(x(t), y(t))dU(t) + \alpha_1(x(t), y(t))dV(t), \qquad x(0) = x_0, \tag{2.1a}$$

$$dy(t) = \frac{1}{\epsilon} g_0(x(t), y(t))dt + \frac{1}{\sqrt{\epsilon}} \beta(x(t), y(t))dV(t), \qquad y(0) = y_0. \tag{2.1b}$$

Later we are going to demonstrate that averaging f_1 and $\alpha_0 \alpha_0^{\mathrm{T}}$ over the invariant measure of the y process with x viewed as fixed, the drift and the diffusion coefficient of the corresponding limiting diffusion process are obtained. The second class is homogenization for SDEs and has the following form

$$\begin{aligned} dx(t) = {} & \left(\frac{1}{\epsilon} f_0(x(t), y(t)) + f_1(x(t), y(t))\right) dt \\ & + \alpha_0(x(t), y(t))dU(t) + \alpha_1(x(t), y(t))dV(t), \qquad x(0) = x_0, \end{aligned} \tag{2.2a}$$

$$dy(t) = \frac{1}{\epsilon^2} g_0(x(t), y(t))dt + \frac{1}{\epsilon} g_1(x(t), y(t))dt + \frac{1}{\epsilon} \beta(x(t), y(t))dV(t), \qquad y(0) = y_0, \tag{2.2b}$$

where $\epsilon << 1$, $x \in \mathcal{X} = \mathbb{T}^l$, $y \in \mathcal{Y} = \mathbb{T}^{d-l}$ (the l and $d-l$ dimensional unit torus respectively), U, V are independent standard Brownian motions of dimension m and n respectively, $f_i : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^l$ ($i \in \{0, 1\}$), $\alpha_0 : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^{l \times n}$, $\alpha_1 : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^{l \times m}$, $g_i(x, y) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^{d-l}$ ($i \in \{0, 1\}$) and $\beta(x, y) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^{(d-l) \times m}$.

As $\epsilon \to 0$ and under the appropriate assumptions on the functions f_0, f_1, g_0 ,g_1, α_0, α_1 and β, the solution $x(t)$ of both systems (2.1) and (2.2) converges in L^p ($p > 1$) in the averaging, and in law in the homogenization case to the solution of a diffusion process of the following form

$$dX(t) = F(X(t))dt + \Sigma(X(t))dW(t), \qquad X(0) = x_0, \tag{2.3}$$

where W is the standard l–dimensional Brownian motion, F and Σ denote the drift and diffusion coefficients respectively and their form can be extracted explicitly.

2.2.1 Limiting equations for multiscale diffusions

In what follows the reader is referred to Pavliotis and Stuart (2008, Chapters 6, 17 and 18) for more details.

Let $\phi^t_\xi(y)$ be the Markov process solving the SDE

$$d(\phi^t_\xi(y)) = g_0(\xi, \phi^t_\xi(y))dt + \beta(\xi, \phi^t_\xi(y))dV(t), \qquad \phi^0_\xi(y) = y_0, \tag{2.4}$$

where g_0 and β as defined above. Then, the generator of this process is

$$\mathcal{L}_0(\xi) = g_0(\xi, y) \cdot \nabla_y + \frac{1}{2}\beta(\xi; y)\beta(\xi; y)^T : \nabla_y \nabla_y. \tag{2.5}$$

Notice that since $\mathcal{Y} = \mathbb{T}^{d-l}$ the operator $\mathcal{L}_0$ and its adjoint $\mathcal{L}_0^*$ are equipped with periodic boundary conditions. The adjoint of the operator $\mathcal{L}_0$ is given by

$$\mathcal{L}_0^* = -\nabla_y g(\xi, y) + \frac{1}{2}\nabla_y \nabla_y \beta(\xi; y)\beta(\xi; y)^T.$$

In the following subsections, we quote the necessary assumptions under which the averaging and homogenization problem converge to a limiting diffusion process and we present the corresponding form of the drift and diffusion coefficients.

2.2.1.1 Averaging

Assumption 2.1.

1. The problem

$$-\mathcal{L}_0^*(\xi)\rho^\infty(y; \xi) = 0 \qquad \& \qquad \int_{\mathcal{Y}} \rho^\infty(y; \xi)dy = 1, \tag{2.6}$$

   has a unique non-negative solution $\rho^\infty(y; \xi) \in L^1(\mathcal{Y})$ for every $\xi \in \mathcal{X}$ and also $\rho^\infty(y; \xi)$ is C^∞ both in y and ξ. This assumption assures us that the fast processes in the systems (2.1b) and (2.2b) are ergodic for each $\xi \in \mathcal{X}$.

2. The functions f_0, f_1, g_0, g_1, α_0, α_1, and β and all derivatives are uniformly bounded in $\mathcal{X} \times \mathcal{Y}$.

3. If $f_0(x, y)$ and all its derivatives with respect to y, ξ are uniformly bounded in $\mathcal{X} \times \mathcal{Y}$ then the same is true for Φ solving the Poisson problem which is defined

by

Definition 2.2.

$$-\mathcal{L}_0\Phi(y;\xi) = f_0(y;\xi), \qquad \int_{\mathcal{Y}} \Phi(y,\xi)\rho^{\infty}(y;\xi)dy = 0. \tag{2.7}$$

Remark 2.3. Provided $\mathcal{Y} = \mathbb{T}^{d-l}$ and $B(\xi, y) := \beta(\xi; y)\beta(\xi; y)^T$ is positive definite all the above assumption hold (see Pavliotis and Stuart (2008, Chapter 6)).

Given Assumption 2.1, define $F : \mathcal{X} \to \mathbb{R}^l$ as

$$F(X) = \int_{\mathcal{Y}} f_0(x,y)\rho^{\infty}(y;x)dy, \tag{2.8}$$

and $\Sigma : \mathcal{X} \to \mathbb{R}^{l\times l}$ as

$$\Sigma(X)\Sigma(X)^T = \int_{\mathcal{Y}} \alpha(x,y)\alpha(x,y)^T\rho^{\infty}(y;x)dy, \tag{2.9}$$

in order for the following theorem to hold.

Theorem 2.4 (Pavliotis and Stuart (2008), Theorem 17.1)**.** *Let $p > 1$. Then, the function $x(t)$ solving the Eq.*(2.1) *converges to $X(t)$ in $L^p(\Omega, C([0,T],\mathcal{X}))$ solving Eq.*(2.3) *where the drift and diffusion coefficients are given by Eq.*(2.8) *and Eq.*(2.9) *respectively.*

2.2.1.2 Homogenization

For the homogenization it is necessary to add, in Assumption 2.1, the following assumption.

Assumption 2.5. The function f_0 averages to zero against the invariant measure of the fast process, namely

$$\int_{\mathcal{Y}} f_0(x,y)\rho^{\infty}(y;x)dy = 0. \tag{2.10}$$

We shall refer to this assumption as the centering condition.

Then, given that the centering condition holds, the Poisson problem (2.7) has a unique solution in $L^p(\mathcal{Y}, \xi)$.

For the homogenization problem, define

$$F(X) = \int_{\mathcal{Y}} f_1(x,y)\rho^{\infty}(y;x)dy, \tag{2.11}$$

and

$$\Sigma(X)\Sigma(X)^T = \Sigma_1(X) + \frac{1}{2}\left(\Sigma_0(X) + \Sigma_0(X)^T\right), \tag{2.12}$$

where

$$\Sigma_1(X) = \int_{\mathcal{Y}} a_0(x,y)a_0(x,y)^T\rho^{\infty}(y;x)dy, \tag{2.13}$$

and

$$\Sigma_0(X) = 2\int_{\mathcal{Y}} f_0(x,y)\otimes\Phi_0(x,y)\rho^{\infty}(y;x)dy. \tag{2.14}$$

Then, under Assumptions 2.1 and 2.5 the following theorem holds.

Theorem 2.6 (Pavliotis and Stuart (2008), Theorem 18.1)**.** *Let $x(t)$ solves Eq.*(2.2) *and $X(t)$ solves Eq.*(2.3) *where the drift and diffusion coefficients are given by Eq.*(2.11) *and Eq.*(2.12) *respectively, then $x(t)$ converges to $X(t)$ in $C([0,T],\mathcal{X})$.*

2.3 Summary

This chapter has introduced a number of key concepts from the theory of multiscale diffusions for which we have reviewed the existing approaches to the problem of parameter estimation in Chapter 1. The performance of these estimators is very satisfactory, however, in order to find the optimal subsampling rate the knowledge of the separation parameter is necessary.

In this thesis, we concentrate on the estimation problem of the diffusion coefficient for the homogenization case. To overcome the issue associated with the separation parameter we propose an estimator that does not require the knowledge of it.

CHAPTER 3

The Extrema Quadratic Variation

We mentioned in Chapter 1 the significance of quadratic variation in the problem of volatility estimation. In this chapter we define formally the quadratic variation of a process, we distinguish its difference with the so–called total quadratic variation and finally we define the extrema quadratic variation (ExtQV) which plays crucial role in our research.

3.1 Types of Quadratic Variation

3.1.1 Quadratic Variation

Let $\pi_{(n)}([0,T]) = \{0 = t_0, t_1, ..., t_n = T\}$ be a partition of size $n+1$ of the $[0,T]$ time interval. Let also, $\text{mesh}(\pi_n([0,T])) := \max_{1\leq i\leq n}(t_i - t_{i-1})$. The quadratic variation of path is defined as follows:

Definition 3.1. *[QV] Let $x(t) : [0,T] \to \mathbb{R}$ be a real–valued continuous path and let $\{x_n(t_i)\}_{i=0}^{n}$ be the piecewise linear approximation of x on a partition $\pi_{(n)}([0,T])$. The quadratic variation (QV) of the path x on the time interval $[0,T]$ is given by*

$$D_2\,(x)_T = \lim_{mesh(\pi_n([0.T]))\to 0} \left(\sum_{t_i\in\pi_{(n)}([0,T])} (\Delta x_n(t_i))^2 \right)^{1/2} . \tag{3.1}$$

3.1.2 Total p–variation

The notion of the total p–variation plays a significant role in the theory of rough paths (see Lyons and Qian (2002); Lyons (1998)).

Definition 3.2 (Total p–variation)**.** *Let $x(t) : [0,T] \to \mathbb{R}$ be a real–valued continuous path. The total p–variation of the path x on the time interval $[0,T]$ is defined by:*

$$D_2^{Total}(x)_T = \sup_{\mathcal{D}([0,T])} \left(\sum_{t_i \in \mathcal{D}([0,T])} |x_{t_i} - x_{t_{i-1}}|^p \right)^{1/p}, \tag{3.2}$$

where $\mathcal{D}([0,T])$ goes through the set of all finite partitions of the interval $[0,T]$.

To visualize the difference between the QV and the total quadratic ($p = 2$) variation consider the Brownian motion process case in $[0,T]$. It is well known that the quadratic variation of the Brownian motion process is bounded and equals to T whereas its total quadratic variation is infinite (see Friz and Victoir (2010, p. 381)).

3.1.3 Extrema Quadratic Variation (ExtQV)

Extrema quadratic variation is the most important notion in our research. It consists our proposed estimator for multiscale processes of bounded variation. Below we give its formal definition.

Definition 3.3. *Let $x(t) : [0,T] \to \mathbb{R}$ be a real–valued continuous path and let $\{x_n(t_i)\}_{i=0}^n$ be the piecewise linear approximation of x. We define the "extrema quadratic variation (ExtQV)" of the path x on the time interval $[0,T]$ as*

$$D_2^{Ext}(x)_T = \lim_{n\to\infty} \left(D_2^{Ext}(x_n)_T \right) = \lim_{n\to\infty} \left(\sum_{\tau_i \in \mathcal{E}_{(n)}([0,T])} (\Delta x_n(\tau_i))^2 \right)^{1/2}, \tag{3.3}$$

where $\mathcal{E}_{(n)}([0,T]) = \{0 = \tau_0, \tau_1, ..., \tau_k = T\}$ is the set of local extremals of $x_n(t)$ and $\Delta x_n(\tau_i) := x_n(\tau_i) - x_n(\tau_{i-1})$, $i = 1, ..., k$.

In other words, let $\pi_n([0,T]) := \{t_0, t_1, ..., t_n\} = \left\{ i\frac{T}{n}, i = 0, ..., n \right\}$ an equally subdivided partition of $[0,T]$ with time step $\delta := T/n$. We say that a point t_i in $\pi_n([0,T])$ is an extremal point and we write $t_i \in \mathcal{E}_{(n)}([0,T])$ if $\Delta x_n(t_i)\Delta x_n(t_{i+1}) = \left(x_{t_i} - x_{t_{i-1}}\right)\left(x_{t_{i+1}} - x_{t_i}\right) < 0$. An illustration of an extremal point can be seen in Figure 3.1 where t_i is an extremal point since $\Delta x_n(t_i) < 0$ and $\Delta x_n(t_{i+1}) > 0$.

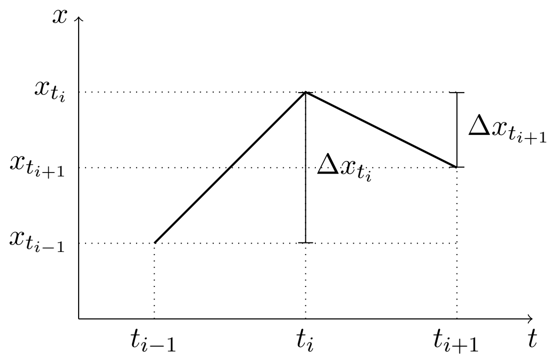

Figure 3.1: Extremal Point Example

Throughout this thesis we denote an extremal points with τ to avoid notation conflicts.

3.2 Computation of (ExtQV) in practice

From the initial partition $\pi_{(n)}([0,T])$, the points that are local extremals for the path are determined to form the corresponding extremal partition, $\mathcal{E}_{(n)}([0,T])$. We call $\{x_n(t_i), t_i \in \pi_{(n)}([0,T])\}$ the original path and $\{x_n(\tau_i), \tau_i \in \mathcal{E}_{(n)}([0,T])\}$ the extremal path. The computation of the quadratic variation requires the consideration of the increments of the original path whereas the extrema quadratic variation the increments of the extremal path. Figure 3.2 illustrates a path observed in 9 time points, $\{t_0,\dots,t_8\}$ and the corresponding extremal path. The black line corresponds to the original path and the red line to the extremal path.

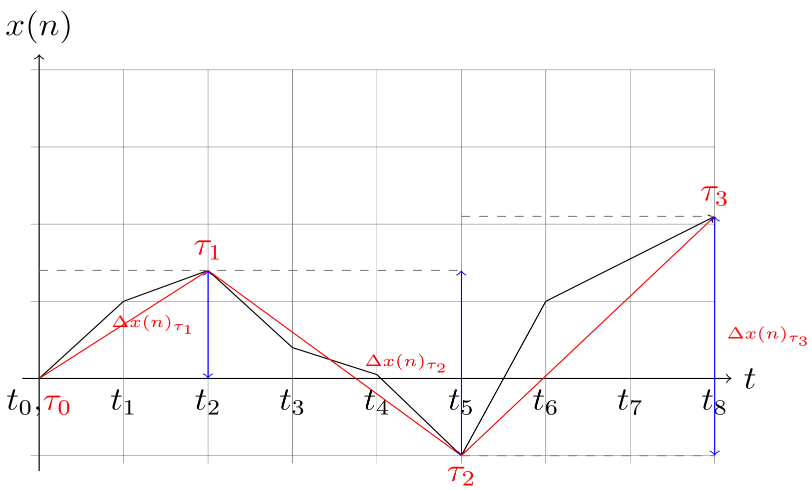

Figure 3.2: The original path (black) and extremal path (red).

In this particular example, the initial partition is $\pi_{(n)}([0,T]) = \{t_0, t_1, ..., t_8\}$ and for the computation of the quadratic variation one has to consider the sum of the squares of all the increments of the original path, namely $(\Delta x(n)_{t_1}, ..., \Delta x(n)_{t_7})$. The corresponding extrema partition is $\mathcal{E}_{(n)}([0,T]) = \{\tau_0, \tau_1, \tau_2, \tau_3\}$ and to compute the extrema quadratic variation the sum of the squared $(\Delta x(n)_{\tau_1}, \Delta x(n)_{\tau_2}, \Delta x(n)_{\tau_3})$ is taken.

Numerically, the procedure below is followed.

Algorithm 1 Extraction of Extremal Path

```
procedure EXTREMALPATH({x_n(t_i), t_i ∈ π_n([0,T])})
    Δx_n(t_i), i ∈ {1,...n}: the increment vector of the path
    temp:={x_n(t_i), t_i ∈ π_n([0,T])}
    for i := 2 to n do
        if Δx_n(t_i)Δx_n(t_{i-1}) > 0 then
            temp(i)='not extremal point'
        end if
    end for
    {x_n(τ_i), τ_i ∈ E_n([0,T])} := temp/('not extremal point')
end procedure
```

The complexity of the above algorithm is of order $\mathcal{O}(n)$ which comes from the for loop and the fact that the computation inside it is of order $\mathcal{O}(1)$.

Having the extrema path, the computation of (ExtQV) simply consists on computing its increments and consecutively take the sum of their squares.

An alternative way to compute the (ExtQV) is to take the sum of squared returns of the original process plus two times the product of those increments such that the consecutive products of the increments between these two are all positive (see Eq.(3.4)). As we shall see in later chapters this way appears to be very useful in the analytic computation of the expectation of (ExtQV).

In what follows define

$$\mathcal{C} := \Big\{ \mathbf{c} := \{c_1, ..., x_k\} \in \mathcal{C} \text{ iff } c_i > 0, \forall i \in \{1, 2, ..., k\} \Big\},$$

a class of vectors whose elements are all positive, and

$$\mathbf{c}^{\mathbf{x}'}_{\mathbf{i,j}} := \Delta x_n(t_i)\Delta x_n(t_j), ..., \Delta x_n(t_i)\Delta x_n(t_{i-1}).$$

Using this notation, the (ExtQV) can be expressed as follows

$$D_2^{\mathrm{Ext}}(x_n)_T^2 = D_2(x_n)_T^2 + 2\sum_{i=2}^{n}\sum_{j=1}^{i-1}\left(\Delta x_n(t_i)\Delta x_n(t_j)\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{i,j}}^{\mathbf{x}'}\right)\right) \tag{3.4}$$

where

$$\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{i,j}}^{\mathbf{x}'}\right) = \begin{cases} 1, & \text{if } \mathbf{c}_{\mathbf{i,j}}^{\mathbf{x}'} \in \mathcal{C}, \\ 0, & \text{if } \mathbf{c}_{\mathbf{i,j}}^{\mathbf{x}'} \notin \mathcal{C}. \end{cases}$$

The following Remark will be very useful in the computation of the expectation of the (ExtQV).

Remark 3.4. For a symmetric and stationary process x the following are true

1. Due to symmetry

$$\mathbb{E}\left[\Delta x_n(t_i)\Delta x_n(t_j)\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{i,j}}^{\mathbf{x}'}\right)\right] = 2\mathbb{E}\left[\Delta x_n(t_i)\Delta x_n(t_j)\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{i,j}}^{\mathbf{x}}\right)\right]$$

   where $\mathbf{c}_{\mathbf{i,j}}^{\mathbf{x}} := \Delta x_n(t_j), ..., \Delta x_n(t_i)$.

2. Due to stationarity

$$\sum_{i=2}^{n}\sum_{j=1}^{i-1}\mathbb{E}\left[\Delta x_n(t_i)\Delta x_n(t_j)\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{i,j}}^{\mathbf{x}}\right)\right] = \sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\Delta x_n(t_1)\Delta x_n(t_k)\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{1,k}}^{\mathbf{x}}\right)\right]$$

Combining these two equations together we get

$$\mathbb{E}\left[D_2^{\mathrm{Ext}}(x_n)_T^2\right] = \mathbb{E}\left[D_2(x_n)_T^2\right] + 4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\Delta x_n(t_1)\Delta x_n(t_k)\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{1,k}}^{\mathbf{x}}\right)\right]. \tag{3.5}$$

In the following chapter we will apply the (ExtQV) to stationary data generated by a simple multiscale model with the fast dynamics described by an Ornstein–Uhlenbeck process. The aim is to show that the (ExtQV) is asymptotically unbiased for the diffusion coefficient of the corresponding limiting equation. To do this we make use of the expression in Eq.(3.5) to compute analytically the expectation of the (ExtQV) when applied to the multiscale data.

CHAPTER 4

Diffusion Coefficient Estimation for a Simple Multiscale Ornstein-Uhlenbeck Process

In this chapter, a simple multiscale Ornstein-Uhlenbeck model is considered for which we prove that our proposed estimator, the extrema quadratic variation (ExtQV), is asymptotically unbiased for the diffusion coefficient of the corresponding coarse–grained model. We also illustrate numerically that the L_2-error is of order ϵ^2. The same model was considered by Papavasiliou (2011) where it was shown that the total p–variation (see Definition 3.2) is asymptotically unbiased for the p–variation of the homogenized SDE and performed with an L_2-error of order ϵ.

4.1 The Model

The model of consideration in this chapter is the following fast/slow system of SDEs

$$dx^\epsilon(t) = \frac{\sigma}{\epsilon} y^\epsilon(t)dt, \qquad x^\epsilon(0) = x_0, \tag{4.1a}$$

$$dy^\epsilon(t) = -\frac{1}{\epsilon^2} y^\epsilon(t)dt + \frac{1}{\epsilon} dW(t), \qquad y^\epsilon(0) = y_0, \tag{4.1b}$$

where W denotes the standard one–dimensional Brownian motion, $\sigma \in \mathbb{R}^+$ is a positive constant and $0 < \epsilon << 1$ denotes a small parameter that controls the scale separation. The fast dynamics are described by an Ornstein-Uhlenbeck process with invariant density the Gaussian distribution with mean zero and variance 1/2.

It is easy to see that the model (4.1) can be equivalently expressed in the following form

$$dx^\epsilon(t) = \sigma \left(dW(t) - \epsilon dy^\epsilon(t) \right), \qquad x^\epsilon(0) = x_0,\ y^\epsilon(0) = y_0. \tag{4.2}$$

Therefore, allowing $\epsilon \to 0$ we deduce that the corresponding limiting homogenized SDE is

$$dX(t) = \sigma dW(t), \qquad X(0) = x_0. \tag{4.3}$$

The objective of this chapter is to propose an asymptotically unbiased estimator for the diffusion coefficient of the coarse–grained model in Eq.(4.3) when using data from Eq.(4.1). It is worth mentioning that the data generated by the model (4.1) are of bounded variation and consequently, their quadratic variation is zero. For this reason, and as discussed in the introductory chapter, we seek for alternative ways to estimate the diffusion coefficient of the corresponding coarse–grained model which in this case is σ. Theorem 4.1 summarizes the main result of this chapter and states that the extrema Quadratic Variation (Definition 3.3) is asymptotically unbiased estimator for σ.

Theorem 4.1. *Let $x^\epsilon(t) : [0,T] \to \mathbb{R}$ be a real–valued path described by Eq.*(4.1)*. Then, as $\epsilon \to 0$, the square of the (ExtQV) is asymptotically unbiased for σ^2 in Eq.*(4.3)*, i.e.*

$$\lim_{\epsilon \to 0} \lim_{n \to \infty} \mathbb{E}\left[D_2^{Ext}(x_n^\epsilon)_T^2\right] = \sigma^2. \tag{4.4}$$

Steps of the proof. To prove the Theorem 4.1 the follow the steps below.

1. Use the expression in Eq.(3.4) and Remark 3.4 to obtain a form of the expectation of (ExtQV) which will allow us to perform explicit computations. In particular, from Eq.(3.4) one has

$$D_2^{\text{Ext}}(x_n^\epsilon)_T^2 = D_2(x_n^\epsilon)_T^2 + 2\sum_{i=2}^{n}\sum_{j=1}^{i-1}\left(\Delta x_n^\epsilon(t_i)\Delta x_n^\epsilon(t_j)\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{j,i}}^{\mathbf{x}^\epsilon}\right)\right), \tag{4.5}$$

   where $\mathcal{C}$ is a class of vectors such that

$$\mathbf{c_{j,i}} := \{c_j, \dots, c_i\} \in \mathcal{C} \text{ iff } c_k > 0, \forall k \in \{j, \dots, i\},$$

$$\mathbf{c}_{\mathbf{j,i}}^{\mathbf{x}^\epsilon} := \{\Delta x_n^\epsilon(t_j)\Delta x_n^\epsilon(t_i), ..., \Delta x_n^\epsilon(t_{i-1})\Delta x_n^\epsilon(t_i)\} \in \mathbb{R}^{i-j},$$

   and

$$\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{j,i}}^{\mathbf{x}^\epsilon}\right) = \begin{cases} 1, & \text{if } \mathbf{c}_{\mathbf{j,i}}^{\mathbf{x}^\epsilon} \in \mathcal{C}, \\ 0, & \text{if } \mathbf{c}_{\mathbf{j,i}}^{\mathbf{x}^\epsilon} \notin \mathcal{C}. \end{cases}$$

   Furthermore, since the process x^ϵ is stationary and symmetric, Remark 3.4

suggests that the expectation of the (ExtQV) takes the following form

$$\mathbb{E}\left[D_2^{\mathrm{Ext}}(x_n^\epsilon)_T^2\right]=\mathbb{E}\left[D_2(x_n^\epsilon)_T^2\right]+4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\Delta x_n^\epsilon(t_1)\Delta x_n^\epsilon(t_k)\mathbf{1}_{\mathcal{C}}\left(\underline{\mathbf{c}}_{\mathbf{1,k}}^{\mathbf{x}^\epsilon}\right)\right], \tag{4.6}$$

where

$$\underline{\mathbf{c}}_{\mathbf{1,k}}^{\mathbf{x}^\epsilon} := \{\Delta x_n^\epsilon(t_1), ..., \Delta x_n^\epsilon(t_k)\} \in \mathbb{R}^k.$$

2. Use the fact that the process x^ϵ is of bounded variation which means that its quadratic variation tends to zero as $n \to \infty$ and thus the first term in Eq.(4.6) is eliminated. Consequently, to prove the result of our theorem it is sufficient to show that

$$\left|4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\Delta x_n^\epsilon(t_1)\Delta x_n^\epsilon(t_k)\mathbf{1}_{\mathcal{C}}\left(\underline{\mathbf{c}}_{\mathbf{1,k}}^{\mathbf{x}^\epsilon}\right)\right]-\sigma^2\right| \to 0, \tag{4.7}$$

as $n \to \infty$ and $\epsilon \to 0$.

3. To work with Eq.(4.7), we express the increments of the x^ϵ process, $\{\Delta x_n^\epsilon(t_i)\}_{i=1}^k$ in terms of the solution to the Ornstein–Uhlenbeck process in Eq.(4.1b). This allow us to use known tools from stochastic calculus to prove analytically Theorem 4.1.

4.2 Analytical Proof of Theorem 4.1

In this section the analytical proof of Theorem 4.1 is presented.

Recall that from Eq.(4.2) the increments of the process x_n^ϵ, Δx_n^ϵ, can be expressed by

$$\Delta x_n^\epsilon(t_i) = \sigma\left(\Delta W_n(t_i) - \epsilon \Delta y_n^\epsilon(t_i)\right). \tag{4.8}$$

A simple application of the Itô formula (see Øksendal (2003)) suggests that the solution of Eq.(4.1b) is given by

$$y_n^\epsilon(t_i) = e^{-\delta/\epsilon^2} y_n^\epsilon(t_{i-1}) + \frac{1}{\epsilon}\int_{t_{i-1}}^{t_i} e^{-\frac{(t_i-u)}{\epsilon^2}}dW(u), \qquad \delta = (T/n). \tag{4.9}$$

The latter allow us to express the Δx_n^ϵ in terms of the process y_n^ϵ, for which we can easily derive its statistics, and a martingale term. Indeed, the increments of the y_n^ϵ

process has the following form

$$\Delta y^{\epsilon}_{t_k}(n) = \left(e^{-\delta/\epsilon^2} - 1\right) y^{\epsilon}_n(t_{k-1}) + \frac{1}{\epsilon}\int_{t_{k-1}}^{t_k} e^{-\frac{(t_k-u)}{\epsilon^2}} dW(u). \tag{4.10}$$

For simplicity define $a := e^{-\delta/\epsilon}$. Then, from equations (4.8) and (4.10) the increments Δx^{ϵ}_n can be expressed by

$$\Delta x^{\epsilon}_n(t_k) = \sigma\left(\epsilon(1-a)y^{\epsilon}_n(t_{k-1}) - \int_{t_{k-1}}^{t_k}\left(e^{-\frac{(t_k-u)}{\epsilon^2}} - 1\right) dW_u\right). \tag{4.11}$$

However, from Eq.(4.6) we are interested on the product $\Delta x^{\epsilon}_n(t_1)\Delta x^{\epsilon}_n(t_k)$ which from Eq.(4.11) is given by

$$\begin{aligned}\Delta x^{\epsilon}_n(t_1)\Delta x^{\epsilon}_n(t_k) \;=\; \sigma^2\Bigg(&\epsilon^2(1-a)^2 y^{\epsilon}_n(t_0)y^{\epsilon}_n(t_{k-1}) \\ &-\epsilon(1-a)y^{\epsilon}_n(t_0)\int_{t_{k-1}}^{t_k}\left(e^{-\frac{(t_k-u)}{\epsilon^2}} - 1\right) dW_u \\ &-\epsilon(1-a)y^{\epsilon}_n(t_{k-1})\int_{t_0}^{t_1}\left(e^{-\frac{(t_1-u)}{\epsilon^2}} - 1\right) dW_u \\ &+\int_{t_0}^{t_1}\left(e^{-\frac{(t_1-u)}{\epsilon^2}} - 1\right) dW_u \int_{t_{k-1}}^{t_k}\left(e^{-\frac{(t_k-u)}{\epsilon^2}} - 1\right) dW_u\Bigg).\end{aligned} \tag{4.12}$$

Also, the condition $\Delta x^{\epsilon}_n(t_i) > 0$, $i \in \{1, ..., k\}$, becomes

$$y^{\epsilon}_n(t_i) > \frac{\int_{t_i}^{t_{i+1}}\left(e^{-\frac{(t_{i+1}-u)}{\epsilon^2}} - 1\right) dW_u}{\epsilon(1-a)} := M_i, \qquad i \in \{0, ..., k-1\}, \tag{4.13}$$

and by defining

$$\mathbf{y}^{\epsilon}_{\mathbf{n}} := \{y^{\epsilon}_n(t_0), ..., y^{\epsilon}_n(t_{k-1})\}\,, \text{ and } \mathbf{M} := \{M_0, ..., M_{k-1}\}, \tag{4.14}$$

Eq.(4.7) given Eq.(4.12) and Eq.(4.13) becomes

$$E \;:=\; \Bigg|\underbrace{4\epsilon^2(1-a)^2\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[y^{\epsilon}_n(t_0)y^{\epsilon}_n(t_{k-1})\mathbf{1}_{\mathcal{C}}\left(\mathbf{y}^{\epsilon}_{\mathbf{n}} - \mathbf{M}\right)\right]}_{E_1} - \sigma^2$$

$$
\underbrace{-4\epsilon^2(1-a)^2 \sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[y_n^\epsilon(t_0)M_{k-1}\mathbf{1}_{\mathcal{C}}\left(\mathbf{y_n^\epsilon}-\mathbf{M}\right)\right]}_{E_2}
$$

$$
\underbrace{-4\epsilon^2(1-a)^2 \sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[y_n^\epsilon(t_{k-1})M_0\mathbf{1}_{\mathcal{C}}\left(\mathbf{y_n^\epsilon}-\mathbf{M}\right)\right]}_{E_3}
$$

$$
\left.\underbrace{+4\epsilon^2(1-a)^2 \sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[M_0M_{k-1}\mathbf{1}_{\mathcal{C}}\left(\mathbf{y_n^\epsilon}-\mathbf{M}\right)\right]}_{E_4}\right|. \quad (4.15)
$$

From triangular inequality one has

$$
E \quad \leq \quad \left|E_1-\sigma^2\right| + |E_2| + |E_3| + |E_4|. \qquad (4.16)
$$

To show that Eq.(4.7) holds it is sufficient to show that each term of Eq.(4.16) goes to zero as $n \to \infty$ and as $\epsilon \to 0$. In what follows we treat each term of Eq.(4.16) separately.

4.2.1 The first term

In this part of the thesis we show that the first term in Eq.(4.16) goes to zero as $n \to \infty$ and then as $\epsilon \to 0$, i.e.,

$$
\lim_{\epsilon\to 0}\lim_{n\to\infty}\left|E_1-\sigma^2\right| = 0. \qquad (4.17)
$$

But, notice that from triangular inequality the following inequality holds

$$
\begin{aligned}
\left|E_1-\sigma^2\right| \quad &= \quad \left|4\sigma^2\epsilon^2(1-a)^2\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[y_n^\epsilon(t_0)y_n^\epsilon(t_{k-1})\mathbf{1}_{\mathcal{C}}\left(\mathbf{y_n^\epsilon}-\mathbf{M}\right)\right]-\sigma^2\right| \\
&\leq \quad \left|4\sigma^2\epsilon^2(1-a)^2\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[y_n^\epsilon(t_0)y_n^\epsilon(t_{k-1})\mathbf{1}_{\mathcal{C}}\left(\mathbf{y_n^\epsilon}\right)\right]-\sigma^2\right| \qquad (4.18a) \\
&\quad + \left|4\sigma^2\epsilon^2(1-a)^2\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[y_n^\epsilon(t_0)y_n^\epsilon(t_{k-1})\mathbf{1}_{\mathcal{C}}\left(|\mathbf{M}|-|\mathbf{y_n^\epsilon}|\right)\right]\right|. \qquad (4.18b)
\end{aligned}
$$

We show the first term (Eq.(4.18a)) tends to zero as $n \to \infty$ and as $\epsilon \to 0$ and that the second term (Eq.(4.18b)) tends to zero as $n \to \infty$.

For the second term, Jensen's inequality (see Lemma A.3) gives

$$\text{Eq.(4.18b)} \quad \leq \quad 4\sigma^2\epsilon^2(1-a)^2 \sum_{k=2}^{n} (n+1-k)\mathbb{E}\left[|y_n^\epsilon(t_0)y_n^\epsilon(t_{k-1})|\, \mathbf{1}_{\mathcal{C}}\left(|\mathbf{M}| - |\mathbf{y}_\mathbf{n}^\epsilon|\right)\right].$$

Furthermore, the condition in the indicator function and the Cauchy–Schwarz inequality (see Lemma A.4) suggest

$$\text{Eq.(4.18b)} \quad \leq \quad 4\sigma^2\epsilon^2(1-a)^2 \sum_{k=2}^{n} (n+1-k) \left(\mathbb{E}\left[|M_0|^2\right] \mathbb{E}\left[|M_{k-1}|^2\right]\right)^{1/2}.$$

In Lemma B.2 we have computed the variance of M to be equal to

$$\mathbb{E}\left[M_i^2\right] = \frac{\frac{1}{3n^3\epsilon^4} + \mathcal{O}\left(\frac{1}{n^4\epsilon^6}\right)}{\epsilon^2(1-a)^2}, \qquad \forall i \in \{0, \ldots k-1\}. \tag{4.19}$$

Given Eq.(4.19), we get

$$\text{Eq.(4.18b)} \quad \leq \quad \frac{4}{3}\sigma^2 \sum_{k=2}^{n} \frac{(n+1-k)}{n^3\epsilon^4} \tag{4.20}$$

which tends to zero as $n \to \infty$.

For Eq.(4.18a), we first consider the expecation

$$E_{11} := \mathbb{E}\left[y_n^\epsilon(t_0)y_n^\epsilon(t_{k-1})\mathbf{1}_{\mathcal{C}}\left(\mathbf{y}_\mathbf{n}^\epsilon\right)\right] \tag{4.21}$$

for which we have

$$\begin{aligned} E_{11} \quad &= \quad \mathbb{E}\left[y_n^\epsilon(t_0)y_n^\epsilon(t_{k-1})\mathbf{1}_{\mathcal{C}}\left(\mathbf{y}_\mathbf{n}^\epsilon\right)\right] \\ &= \quad \mathbb{E}\left[y_n^\epsilon(t_0)y_n^\epsilon(t_{k-1}) \prod_{i=0}^{k-1} \mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_i)\right)\right] \end{aligned} \tag{4.22}$$

Applying Tower property (see Lemma A.1), Eq.(4.22) is equivalent to

$$E_{11} \quad = \quad \mathbb{E}\left[\mathbb{E}\left[y_n^\epsilon(t_0)y_n^\epsilon(t_{k-1}) \prod_{i=0}^{k-1} \mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_i)\right) \middle| y_n^\epsilon(t_0), \ldots, y_n^\epsilon(t_{k-2})\right]\right],$$

and since the process y_n^ϵ is a Markov process we get

$$E_{11} = \mathbb{E}\left[\mathbb{E}\left[y_n^\epsilon(t_0)y_n^\epsilon(t_{k-1})\prod_{i=0}^{k-1}\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_i)\right)\middle| y_n^\epsilon(t_{k-2})\right]\right].$$

Finally, taking out what is known (TOWIK) property (see Lemma A.2) suggests

$$E_{11} = \mathbb{E}\left[y_n^\epsilon(t_0)\prod_{i=0}^{k-2}\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_i)\right)\mathbb{E}\left[y_n^\epsilon(t_{k-1})\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_{k-1})\right)\middle| y_n^\epsilon(t_{k-2})\right]\right]. \quad (4.23)$$

In Lemma 4.2 we prove an inequality for the inner expectation in Eq.(4.23) which will be very useful in the computation of the desired result.

Lemma 4.2. *Let* $K_{t_i} := \frac{1}{\epsilon}\int_{t_i}^{t_{i+1}} e^{-\frac{t_{i+1}-u}{\epsilon^2}}dW_u$ *which is normally distributed with zero mean and variance* σ_K^2. *Let also* $\phi_K(\cdot)$, $\Phi_K(\cdot)$ *be the pdf and cdf of* K *respectively. Then, Eq.*(4.9) *is equal to*

$$y_n^\epsilon(t_i) = ay_n^\epsilon(t_{i-1}) + K_{t_{i-1}}, \qquad \delta = (T/n), \quad (4.24)$$

and

$$\begin{aligned} IE_{11} : &= \mathbb{E}\left[y_n^\epsilon(t_{k-1})\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_{k-1})\right)\middle| y_n^\epsilon(t_{k-2})\right] \\ &\leq \left(ay_n^\epsilon(t_{k-2}) + \sigma_K^2\phi_K\left(-ay_n^\epsilon(t_{k-2})\right)\right)\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_{k-2})\right). \end{aligned} \quad (4.25)$$

Proof.

$$\begin{aligned} \mathrm{IE}_{11} &= \mathbb{E}\left[y_n^\epsilon(t_{k-1})\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_{k-1})\right)\middle| y_n^\epsilon(t_{k-2})\right] \\ &\overset{\text{Eq.(4.24)}}{=} \mathbb{E}\left[ay_n^\epsilon(t_{k-2})\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_{k-1})\right)\middle| y_n^\epsilon(t_{k-2})\right] \\ &\qquad +\mathbb{E}\left[K_{t_{k-2}}\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_{k-1})\right)\middle| y_n^\epsilon(t_{k-2})\right] \\ &\overset{\text{(TOWIK)}}{=} ay_n^\epsilon(t_{k-2})\mathbb{E}\left[\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_{k-1})\right)\middle| y_n^\epsilon(t_{k-2})\right] \\ &\qquad +\mathbb{E}\left[K_{t_{k-2}}\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_{k-1})\right)\middle| y_n^\epsilon(t_{k-2})\right] \\ &\overset{\text{Eq.(4.24)}}{=} ay_n^\epsilon(t_{k-2})\mathbb{P}\left[\left(K_{t_{k-2}} > -ay_n^\epsilon(t_{k-2})\right)\middle| y_n^\epsilon(t_{k-2})\right] \\ &\qquad +\mathbb{E}\left[K_{t_{k-2}}\mathbf{1}_{\mathcal{C}}\left(K_{t_{k-2}} + ay_n^\epsilon(t_{k-2})\right)\middle| y_n^\epsilon(t_{k-2})\right] \end{aligned}$$

$$
\begin{aligned}
&= a y_n^\epsilon(t_{k-2}) - a y_n^\epsilon(t_{k-2}) \mathbb{P}\left[\left(K_{t_{k-2}} \leq -a y_n^\epsilon(t_{k-2}) \right) \middle| y_n^\epsilon(t_{k-2}) \right] \\
&\quad + \mathbb{E}\left[K_{t_{k-2}} \mathbf{1}_{\mathcal{C}} \left(K_{t_{k-2}} + a y_n^\epsilon(t_{k-2}) \right) \middle| y_n^\epsilon(t_{k-2}) \right] \\
&= a y_n^\epsilon(t_{k-2}) - a y_n^\epsilon(t_{k-2}) \Phi_K \left(-a y_n^\epsilon(t_{k-2}) \right) + \sigma_K^2 \phi_K \left(-a y_n^\epsilon(t_{k-2}) \right).
\end{aligned}
$$

From Lemma B.4 we get

$$
\begin{aligned}
\mathrm{IE}_{11} &\leq a y_n^\epsilon(t_{k-2}) + \left(\sigma_K^2 \phi_K \left(-a y_n^\epsilon(t_{k-2}) \right) \right) \mathbf{1}_{\mathcal{C}} \left(y_n^\epsilon(t_{k-2}) \right) \\
&\quad + \left(\sigma_K^2 \phi_K \left(-a y_n^\epsilon(t_{k-2}) \right) - a y_n^\epsilon(t_{k-2}) \right) \mathbf{1}_{\mathcal{C}} \left(-y_n^\epsilon(t_{k-2}) \right) \\
&= \left(a y_n^\epsilon(t_{k-2}) + \sigma_K^2 \phi_K \left(-a y_n^\epsilon(t_{k-2}) \right) \right) \mathbf{1}_{\mathcal{C}} \left(y_n^\epsilon(t_{k-2}) \right) \qquad (4.26)
\end{aligned}
$$

as required. □

Now, Eq.(4.23) given the result in Lemma 4.2 becomes

$$
\begin{aligned}
E_{11} &\leq a \mathbb{E}\left[y_n^\epsilon(t_0) y_n^\epsilon(t_{k-2}) \prod_{i=0}^{k-2} \mathbf{1}_{\mathcal{C}} \left(y_n^\epsilon(t_i) \right) \right] \qquad (4.27a) \\
&\quad + \sigma_K^2 \mathbb{E}\left[y_n^\epsilon(t_0) \phi_K \left(-a y_n^\epsilon(t_{k-2}) \right) \prod_{i=0}^{k-2} \mathbf{1}_{\mathcal{C}} \left(y_n^\epsilon(t_i) \right) \right]. \qquad (4.27b)
\end{aligned}
$$

Applying the same procedure on Eq.(4.27a) one has

$$
\begin{aligned}
E_{11} &\leq a^2 \mathbb{E}\left[y_n^\epsilon(t_0) y_n^\epsilon(t_{k-3}) \prod_{i=0}^{k-3} \mathbf{1}_{\mathcal{C}} \left(y_n^\epsilon(t_i) \right) \right] \qquad (4.28a) \\
&\quad + a \sigma_K^2 \mathbb{E}\left[y_n^\epsilon(t_0) \phi_K \left(-a y_n^\epsilon(t_{k-3}) \right) \prod_{i=0}^{k-3} \mathbf{1}_{\mathcal{C}} \left(y_n^\epsilon(t_i) \right) \right]. \\
&\qquad (4.28b)
\end{aligned}
$$

Combining Eq.(4.27) and Eq.(4.28) gives

$$
\begin{aligned}
E_{11} &\leq a^2 \mathbb{E}\left[y_n^\epsilon(t_0) y_n^\epsilon(t_{k-3}) \prod_{i=0}^{k-3} \mathbf{1}_{\mathcal{C}} \left(y_n^\epsilon(t_i) \right) \right] \qquad (4.29) \\
&\quad + a \sigma_K^2 \mathbb{E}\left[y_n^\epsilon(t_0) \phi_K \left(-a y_n^\epsilon(t_{k-3}) \right) \prod_{i=0}^{k-3} \mathbf{1}_{\mathcal{C}} \left(y_n^\epsilon(t_i) \right) \right] \\
&\quad + \sigma_K^2 \mathbb{E}\left[y_n^\epsilon(t_0) \phi_K \left(-a y_n^\epsilon(t_{k-2}) \right) \prod_{i=0}^{k-2} \mathbf{1}_{\mathcal{C}} \left(y_n^\epsilon(t_i) \right) \right], \qquad (4.30)
\end{aligned}
$$

and after $k-1$ times

$$\begin{aligned} E_{11} \quad \leq \quad & a^{k-1}\mathbb{E}\left[y_n^\epsilon(t_0)^2\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_0)\right)\right] \\ & +\sigma_K^2\sum_{j=0}^{k-2}a^j\mathbb{E}\left[y_n^\epsilon(t_0)\phi_K\left(-ay_n^\epsilon(t_{k-2-j})\right)\prod_{i=0}^{k-2-j}\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_i)\right)\right]. \end{aligned} \tag{4.31}$$

Finally, Eq.(4.18a) given Eq.(4.31) takes the following form

$$\text{Eq.(4.18a)} \quad \leq \quad \left|4\sigma^2\epsilon^2(1-a)^2\sum_{k=2}^{n}(n+1-k)a^{k-1}\mathbb{E}\left[y_n^\epsilon(t_0)^2\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_0)\right)\right]-\sigma^2\right| \tag{4.32a}$$

$$\begin{aligned} +\Bigg|4\sigma^2\epsilon^2(1-a)^2\sigma_K^2\sum_{k=2}^{n}(n+1-k)\sum_{j=0}^{k-2}a^j\times \\ \mathbb{E}\left[y_n^\epsilon(t_0)\phi_K\left(-ay_n^\epsilon(t_{k-2-j})\right)\prod_{i=0}^{k-2-j}\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_i)\right)\right]\Bigg|. \end{aligned} \tag{4.32b}$$

Since $y_n^\epsilon(t_0)\sim\mathcal{N}\left(0,\frac{1}{2}\right)$ we obtain $\mathbb{E}\left[y_n^\epsilon(t_0)^2\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_0)\right)\right]=\frac{1}{4}$ and thus

$$\begin{aligned} & 4\sigma^2\epsilon^2(1-a)^2\sum_{k=2}^{n}(n+1-k)a^{k-1}\mathbb{E}\left[y_n^\epsilon(t_0)^2\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_0)\right)\right] \\ & \qquad =\sigma^2\epsilon^2(1-a)^2\sum_{k=2}^{n}(n+1-k)a^{k-1} \\ & \qquad \to\sigma^2\left(1+\epsilon^2\left(e^{-1/\epsilon^2}-1\right)\right)\ \text{ as } n\to\infty \text{ (see Lemma B.1)} \\ & \qquad \to\sigma^2 \text{ as } \epsilon\to 0\ . \end{aligned}$$

Therefore, for Eq.(4.32a) we get

$$\left|4\sigma^2\epsilon^2(1-a)^2\sum_{k=2}^{n}(n+1-k)a^{k-1}\mathbb{E}\left[y_n^\epsilon(t_0)^2\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_0)\right)\right]-\sigma^2\right|\to 0 \tag{4.33}$$

as $n\to\infty$ and as $\epsilon\to 0$.

For Eq.(4.32b), since $\sigma_K^2=\dfrac{1-a^2}{2}$ (see Lemma B.3) and the $K_i\overset{i.i.d}{\sim}\mathcal{N}(0,\sigma_K^2)$ we get

$$\begin{aligned}
\text{Eq.(4.32b)} =& \Bigg| 2\sigma^2\epsilon^2(1-a)^2(1-a^2)\sum_{k=2}^{n}(n+1-k)\sum_{j=0}^{k-2}a^j\times \\
& \mathbb{E}\left[y_n^\epsilon(t_0)\phi_K\left(-ay_n^\epsilon(t_{k-2-j})\right)\prod_{i=0}^{k-2-j}\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_i)\right)\right]\Bigg| \\
=& \Bigg| \frac{2\sigma^2\epsilon^2}{\sqrt{\pi}}(1-a)^2(1-a^2)^{1/2}\sum_{k=2}^{n}(n+1-k)\sum_{j=0}^{k-2}a^j\times \\
& \mathbb{E}\left[y_n^\epsilon(t_0)e^{-\frac{\left(ay_n^\epsilon(t_{k-2-j})\right)^2}{2\sigma_K^2}}\prod_{i=0}^{k-2-j}\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_i)\right)\right]\Bigg| && (4.34) \\
\leq& \frac{2\sigma^2\epsilon^2}{\sqrt{\pi}}(1-a)^2(1-a^2)^{1/2}\sum_{k=2}^{n}(n+1-k)\sum_{j=0}^{k-2}a^j\times \\
& \mathbb{E}\left[\left|y_n^\epsilon(t_0)e^{-\frac{\left(ay_n^\epsilon(t_{k-2-j})\right)^2}{2\sigma_K^2}}\prod_{i=0}^{k-2-j}\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_i)\right)\right|\right] && (4.35) \\
\leq& \frac{2\sigma^2\epsilon^2}{\sqrt{\pi}}(1-a)^2(1-a^2)^{1/2}\sum_{k=2}^{n}(n+1-k)\sum_{j=0}^{k-2}a^j\times \\
& \mathbb{E}\left[y_n^\epsilon(t_0)^2\right]^{1/2}\mathbb{E}\left[e^{-\frac{\left(ay_n^\epsilon(t_{k-2-j})\right)^2}{\sigma_K^2}}\right]^{1/2}\mathbb{E}\left[\left(\prod_{i=0}^{k-2-j}\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_i)\right)\right)^2\right]^{1/2}. && (4.36)
\end{aligned}$$

But,

$$\mathbb{E}\left[y_n^\epsilon(t_0)^2\right] = \frac{1}{2}, \quad \mathbb{E}\left[e^{-\frac{\left(ay_n^\epsilon(t_{k-2-j})\right)^2}{\sigma_K^2}}\right] = \left(\frac{1-a^2}{1+a^2}\right)^{1/2}, \tag{4.37}$$

and

$$\begin{aligned}
\mathbb{E}\left[\left(\prod_{i=0}^{k-2-j}\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_i)\right)\right)^2\right] &= \mathbb{E}\left[\prod_{i=0}^{k-2-j}\mathbf{1}_{\mathcal{C}}\left(y_n^\epsilon(t_i)\right)\right] \\
&= \mathbb{P}\left[y_n^\epsilon(t_0)>0\right]\mathbb{P}\left[y_n^\epsilon(t_1)>0\middle|y_n^\epsilon(t_0)>0\right] \\
&\quad\times\mathbb{P}\left[y_n^\epsilon(t_2)>0\middle|y_n^\epsilon(t_1)>0, y_n^\epsilon(t_0)>0\right] \\
&\quad\ddots
\end{aligned}$$

$$\times \mathbb{P}\left[y_n^\epsilon(t_{k-2-j}) > 0 \,\middle|\, \cap_{i=0}^{k-3-j} y_n^\epsilon(t_i) > 0\right] \tag{4.38}$$

$$= \mathbb{P}\left[y_n^\epsilon(t_0) > 0\right] \prod_{i=1}^{k-2-j} \mathbb{P}\left[y_n^\epsilon(t_i) > 0 \,\middle|\, y_n^\epsilon(t_{i-1}) > 0\right] \tag{4.39}$$

$$= \frac{1}{2}\left(\mathbb{P}\left[y_n^\epsilon(t_1) > 0 \,\middle|\, y_n^\epsilon(t_0) > 0\right]\right)^{k-2-j}. \tag{4.40}$$

For the equality in Eq.(4.38) we have applied the general multiplicative rule for conditional probabilities, for Eq.(4.39) we have used the Markov property and finally in equality Eq.(4.40) we have used the fact that the process y_n^ϵ is stationary. In Lemma B.5 we have computed the probability in Eq.(4.40) to be equal to

$$\rho_{n,\epsilon} := \mathbb{P}\left[y_n^\epsilon(t_1) > 0 \,\middle|\, y_n^\epsilon(t_0) > 0\right] = 1 - \frac{\arctan \frac{(1-a^2)}{a}}{\pi}. \tag{4.41}$$

Thus, Eq.(4.32b) given Eq.(4.36), Eq.(4.37), Eq.(4.40) and Eq.(4.41) becomes

$$\text{Eq.(4.32b)} \leq \frac{2\sigma^2\epsilon^2}{\sqrt{2\pi}}(1-a)^2 \frac{(1-a^2)^{3/4}}{(1+a^2)^{1/4}} \sum_{k=2}^{n}(n+1-k)\sum_{j=0}^{k-2} a^j \rho_{n,\epsilon}^{\frac{k-2-j}{2}}.$$

In Lemma B.6 we have shown that the above sum tends to zero as $n \to \infty$ which concludes the proof for the first term, i.e., Eq.(4.17) holds. In what follows we show that the rest of the terms in Eq.(4.15) also tend to zero as $n \to \infty$. As we shall see the situation is much simpler for these terms.

4.2.2 The second the third term

For the second term, $|E_2|$, we apply Jensen's and Cauchy–Schwarz and together with the fact that $\mathbf{1}_{\mathcal{C}}\left(\mathbf{y_n^\epsilon} - \mathbf{M}\right) \leq 1$ we obtain

$$\begin{aligned}
|E_2| &= \left| 4\epsilon^2(1-a)^2 \sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[y_n^\epsilon(t_0) M_{t_{k-1}} \mathbf{1}_{\mathcal{C}}\left(\mathbf{y_n^\epsilon} - \mathbf{M}\right)\right] \right| \\
&\leq 4\epsilon^2(1-a)^2 \sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\left|y_n^\epsilon(t_0) M_{t_{k-1}}\right|\right] \\
&\leq 4\epsilon^2(1-a)^2 \sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[y_n^\epsilon(t_0)^2\right]^{1/2} \mathbb{E}\left[M_{t_{k-1}}^2\right]^{1/2}
\end{aligned}$$

$$= \frac{4}{\sqrt{6}}(1-a)\sum_{k=2}^{n}\frac{(n+1-k)}{n^{3/2}\epsilon}, \text{ as } n \to \infty \tag{4.42}$$

with the latter equality coming from the result in Lemma B.2.

Exactly the same approach can be followed to show that $|E_3| \to 0$ as $n \to 0$.

4.2.3 The fourth term

Finally, for the fourth term, $|E_4|$, similar arguments with before are applied to get

$$\begin{aligned}
|E_4| &= \left|4\epsilon^2(1-a)^2\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[M_0 M_{k-1}\mathbf{1}_{\mathcal{C}}\left(\mathbf{y}_{\mathbf{n}}^{\epsilon}-\mathbf{M}\right)\right]\right| \\
&\leq 4\epsilon^2(1-a)^2\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[|M_0 M_{k-1}|\right] \\
&\leq \epsilon^2(1-a)^2\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[M_{t_0}^2\right]^{1/2}\mathbb{E}\left[M_{t_{k-1}}^2\right]^{1/2} \\
&= \frac{4}{3\epsilon^4}\sum_{k=2}^{n}\frac{(n+1-k)}{n^3} \to 0, \text{ as } n \to \infty
\end{aligned} \tag{4.43}$$

with the latter equality coming from the result in Lemma B.2.

The results in the subsections 4.2.1, 4.2.3 and 4.2.2 complete the proof of Theorem 4.1.

In the following section we shall illustrate the validity of our theoretical arguments numerically.

4.3 Numerical Results

In this section, we examine numerically the performance of the (ExtQV) when it is applied to data generated by the model (4.1). First, we perform computations for its expectation to examine if it is indeed unbiased as we have shown in Section 4.2. Secondly, we examine its consistency by computing the L_2-error.

4.3.1 Unbiasedness of the Exrema Quadratic Variation

Here, a numerical study is presented to explore the unbiasedness of our proposed estimator. We also examine how the choice of the parameter ϵ, the step size $\delta = T/n$ and σ affects the accuracy of the (ExtQV). Unless stated otherwise, $T = 1$. In this way we control the step size by choosing the values of n.

We generate 1000 realizations of the path x_n^ϵ of size n using Euler–Maruyama scheme. For each realization we evaluate the (ExtQV) using the Algorithm 1 which it is described in Subsection 3.2. For the expectation we take the average of these values. A rough estimate of the numerical error for the computation of the expectation is of $\mathcal{O}(10^{-2})$.

$\mathbb{E}\left[D_2^{\text{Ext}}(x_n^\epsilon)_T^2\right]$	ϵ			
	0.05	0.10	0.15	0.20
$n = 10^3$	1.4971	1.1956	1.0706	1.0556
$n = 10^4$	1.1317	1.0569	1.0327	0.9639
$n = 10^5$	1.0400	0.9997	1.0003	0.9682
$n = 10^6$	1.0085	0.9861	0.9599	0.9447
$n = 10^7$	0.9908	0.9904	0.9665	0.9515
Theoretical Value	0.9975	0.9900	0.9775	0.9600

Table 4.1: Expectation of the (ExtQV) for different n's and ϵ's and for $\sigma = 1$.

Table 4.1 presents the values of the expectation of the (ExtQV) for four different values of $\epsilon = (0.05, 0.10, 0.15, 0.20)$, five values of $n = \left(10^3, 10^4, 10^5, 10^6, 10^7\right)$ and for $\sigma = 1$. The last line of the table corresponds to the theoretical value of the (ExtQV) as $n \to \infty$ which is given by

$$\lim_{n\to\infty} \mathbb{E}\left[D_2^{\text{Ext}}(x_n^\epsilon)_T^2\right] = \sigma^2 \left(1 + \epsilon^2 \left(1 - e^{-1/\epsilon^2}\right)\right).$$

Figure 4.1 presents the value of the expectation of the (ExtQV) for a range of ϵ values in the interval $[0.01, 0.20]$ and the five values of n. This figure emphasizes the fact that a sufficient large of sample size n should be considered in order the quantity $\frac{1}{n\epsilon^2} << 1$. This ensures that the error due to the discretization is negligible.
Indeed, we observe that the smaller the value of ϵ is, the larger the sample size should be in order for the (ExtQV) to perform as expected. For example, considering both Table 4.1 and Figure 4.1, when ϵ=0.20 the (ExtQV) converges much faster to the expected value than for $\epsilon = 0.05$.

Now we are going to fix n and investigate the effect of σ value. Figures 4.2 and 4.3

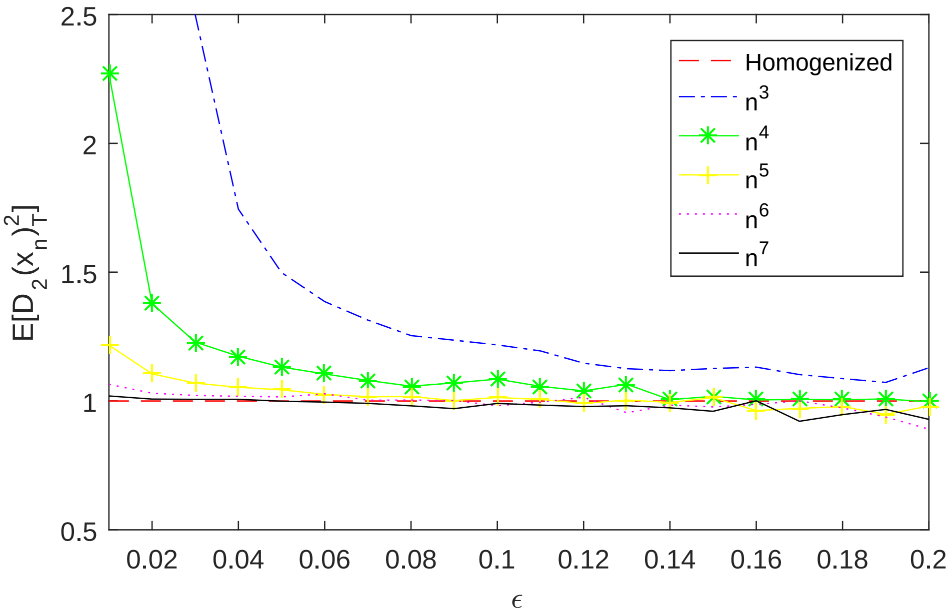

Figure 4.1: Expectation of the (ExtQV) for different n's and ϵ's and for $\sigma = 1$.

show the behavior of the (ExtQV) with respect to the scale separation parameter ϵ for $\sigma = 1$ and $\sigma = 2$, respectively. For a sufficient large sample size of our data (in this case $n = 10^6$), as the value of separation parameter is getting smaller then the expectation of the (ExtQV) is getting closer to the square of the homogenized diffusion coefficient. For this reason in Figure 4.2 is going to 1 and in Figure 4.3 is going to 4.

Finally, in Figure 4.4, we fix ϵ and n and we examine the behavior of our estimator for different values of σ. The heuristics in Figure 4.4 suggest that the expectation of the (ExtQV) achieves the value of the homogenized coefficient for any value of σ.

4.3.2 Consistency of the Extrema Quadratic Variation

The consistency of the (ExtQV) is examined via the L_2–error, i.e.

$$\mathbb{E}\left[\left(\mathbb{E}\left(D_2^{\text{Ext}}(x_n^\epsilon)_T\right)^2 - \sigma^2\right)^2\right].$$

Table 4.2 shows the L_2–error for different n's, ϵ's and for fixed $\sigma = 1$. In the Appendix B can be found similar tables that correspond to $\sigma = (2, 3, 4)$, see Tables B.1 and B.2. Based on these results and Figure 4.5 we conclude that as $n \to \infty$ our

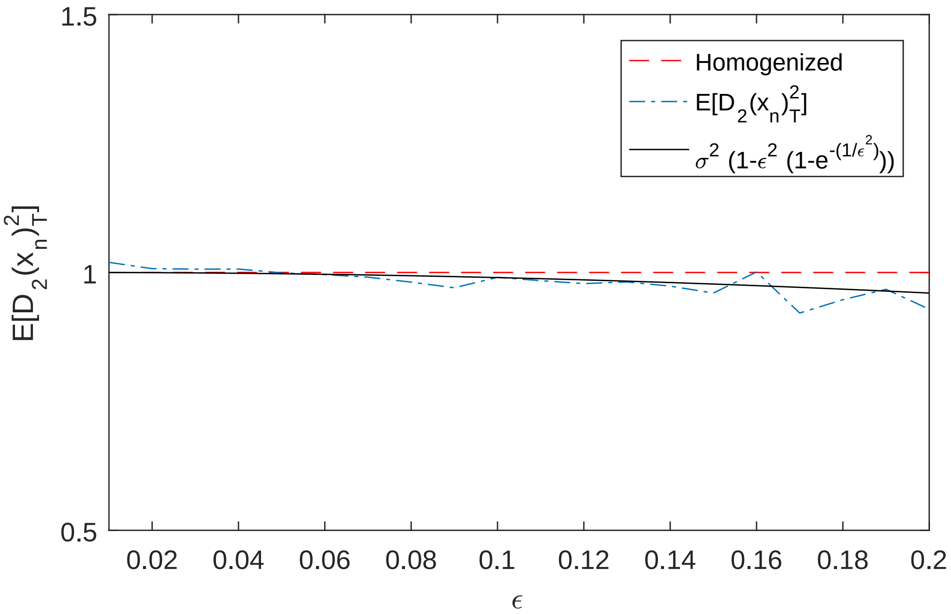

Figure 4.2: Expectation of the (ExtQV) for different ϵ's and for fixed $n = 10^6$ and $\sigma = 1$.

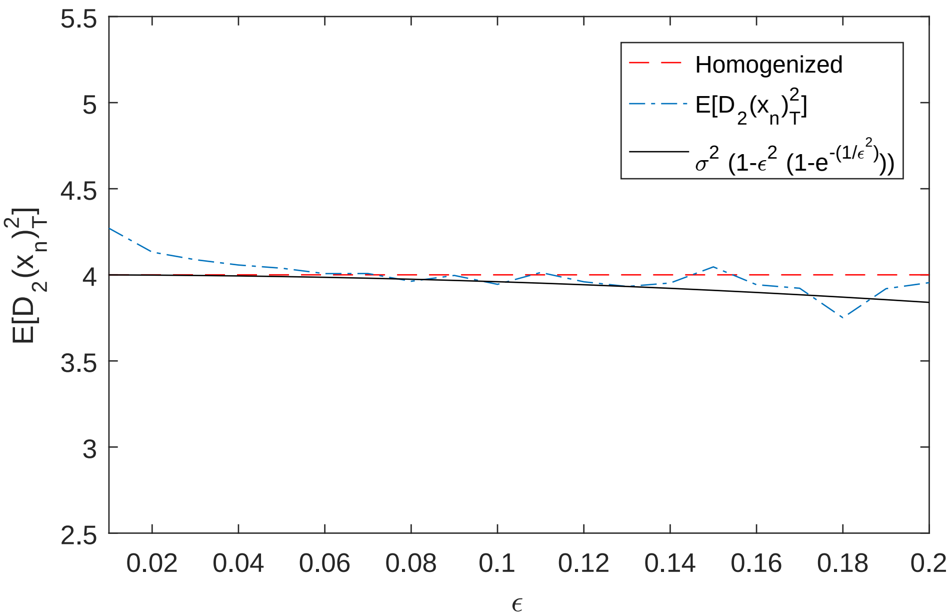

Figure 4.3: Expectation of the (ExtQV) for different ϵ's and for fixed $n = 10^6$ and $\sigma = 2$.

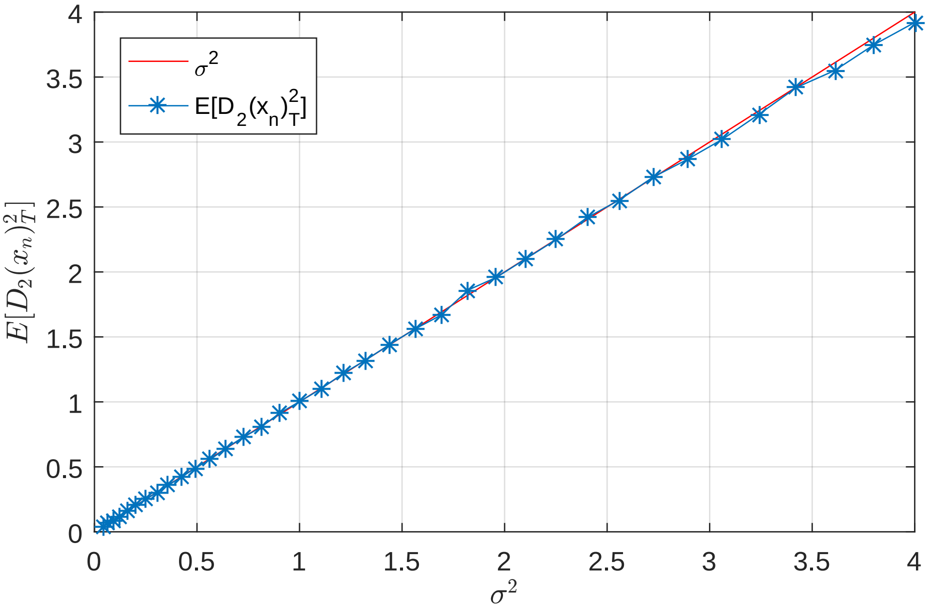

Figure 4.4: Expectation of the (ExtQV) for different σ's and for fixed $n = 10^6$ and $\epsilon = 0.10$.

$\mathbb{E}\left[\left(D_2^{\text{Ext}}(x_n^\epsilon)_T^2 - \sigma^2\right)^2\right]$	ϵ			
	0.05	0.10	0.15	0.20
$n = 10^3$	0.2985	0.1785	0.1792	0.3638
$n = 10^4$	0.0476	0.1250	0.2650	0.3548
$n = 10^5$	0.0306	0.1154	0.2380	0.3800
$n = 10^6$	0.0273	0.10	0.1986	0.3874
$n = 10^7$	0.0261	0.1008	0.1992	0.3824

Table 4.2: L_2–error of the (ExtQV) for different n's and ϵ's and for $\sigma = 1$.

estimator performs with an L_2–error of order ϵ^2.

Figure 4.5 represents the log–log plot between the ratio of the L_2–error corresponding to $k\epsilon$ and the L_2–error corresponding to ϵ with respect to $\log(k\epsilon)$ where $k = 1, \ldots, 20$. The behavior of the log–log plot is linear with slop roughly equal to 2 (red line). For example this means that the L_2 corresponding to $2\epsilon = 0.02$ is 4 times the L_2–error corresponding to ϵ times a quantity of order ϵ^2 (for this example $4\epsilon^2$).

4.4 Summary

As it was mentioned in the beginning of this chapter, the diffusion coefficient estimation problem for the model introduced in Section 4.1 was also considered in

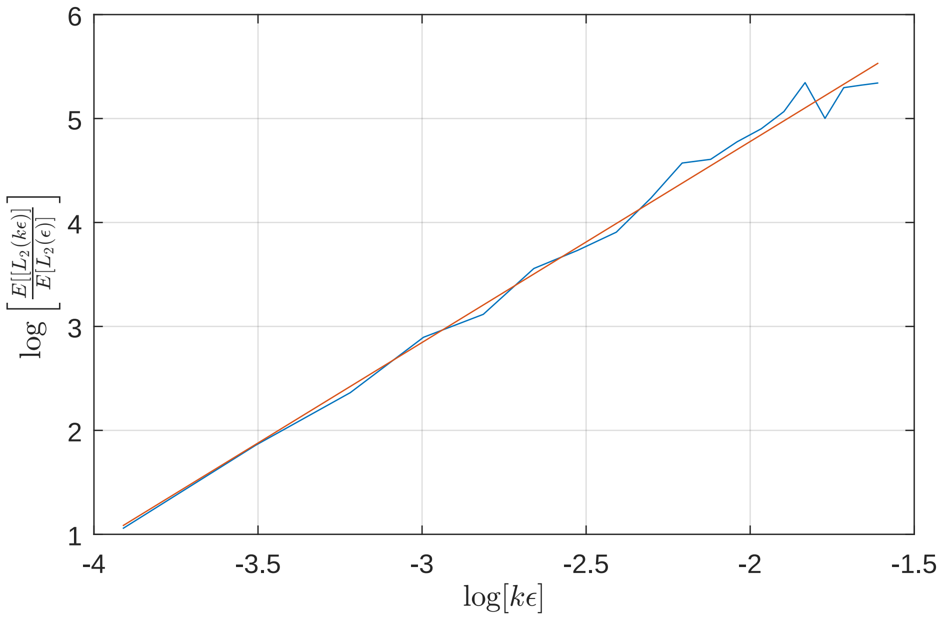

Figure 4.5: Log–log plot between the ration of the L_2–error corresponding to $k\epsilon$ and L_2–error corresponding to ϵ with respect to $\log(k\epsilon)$.

Papavasiliou (2011). The author proved that the total quadratic variation (Definition 3.2 for $p = 2$) consists an unbiased estimator for the diffusion coefficient of the limiting diffusion process with an L_2–error of order ϵ. Here, we showed that for the same model the (ExtQV) is asymptotically unbiased for our parameter of interest. Also, numerical evidence suggests that the (ExtQV) performs with an L_2–error of order ϵ^2.

CHAPTER 5

General Bounded Variation Model

The objective of this chapter is to examine if the proposed estimator can be used for more general models with zero quadratic variation. We start by introducing the general form of the multiscale models that will be considered in this chapter. Then, the analytical proof that the (ExtQV) is asymptotically unbiased for this class of models is presented. Finally, we extend our results when the corresponding homogenized diffusion contains also a drift term.

5.1 The General Model

In the general case, fast/slow systems of SDEs of the following form are considered

$$dx^\epsilon(t) = \frac{1}{\epsilon} f(y^\epsilon(t))dt, \qquad x^\epsilon(0) = x_0^\epsilon, \tag{5.1a}$$

$$dy^\epsilon(t) = \frac{1}{\epsilon^2} g(y^\epsilon(t))dt + \frac{\beta(y^\epsilon)}{\epsilon} dV(t), \qquad y^\epsilon(0) = y_0^\epsilon, \tag{5.1b}$$

where $(x^\epsilon, y^\epsilon) \in \mathcal{X} \times \mathcal{Y} = \mathbb{T} \times \mathbb{T}$, ($\mathbb{T}$ is the unit torus) V is the standard one–dimensional Browinan motion and the functions f, g and β and all their derivatives are continuous, smooth and uniformly bounded on the torus. Under these assumption the generator of the y^ϵ process, $\mathcal{L}_0$, is a bounded operator with respect to the L^∞ norm on the torus. Furthermore, we assume that its inverse, $\mathcal{L}_0^{-1}$, exist and is also bounded with respect to the L^∞ norm.

The theory in Chapter 2 suggests that as $\epsilon \to 0$, the process x^ϵ converges weakly to the process X solving the following SDE

$$dX(t) = \sigma dW(t), \qquad X(0) = x_0, \tag{5.2}$$

where W is the standard Brownian motion and is independent of V. The diffusion coefficient σ, which for this class of models is constant, is given by

$$\sigma^2 = 2\int_{\mathcal{Y}} f(y^\epsilon)\Phi(y^\epsilon)\rho^\infty(y^\epsilon)dy^\epsilon = 2\mathbb{E}_{y^\epsilon}\left[f(y^\epsilon)\Phi(y^\epsilon)\right], \tag{5.3}$$

where the expectation is with respect to the invariant density of the y^ϵ process. The function $\Phi(\cdot)$ solves the Poisson problem

$$\begin{aligned} \left(\mathcal{L}_0\Phi\right)(y^\epsilon) &= -f(y^\epsilon), \\ \int_{\mathcal{Y}} \Phi(y^\epsilon)\rho^\infty(y^\epsilon)dy^\epsilon &= 0, \\ \Phi(y^\epsilon) \quad &\text{is} \quad \text{periodic on } \mathcal{Y}. \end{aligned} \tag{5.4}$$

Remark 5.1. Since y^ϵ is a Markov process, its corresponding generator, $\mathcal{L}_0$, is a second order elliptic operator and Backward Kolmogorov Equation (BKE) becomes an initial value problem for parabolic PDEs (see Pavliotis (2014)).

By Fredholm alternative for elliptic PDEs with periodic boundary conditions, Eq.(5.4) has a unique centered solution, see Pavliotis and Stuart (2008, Chapter 6).

Remark 5.2 (Lemma 18.3 in Pavliotis and Stuart (2008)). The function Φ and all its derivatives are smooth and uniformly bounded.

A necessary assumption for the multiscale model (5.1) to produce a sensible limit as $\epsilon \to 0$ is the Assumption 2.5, i.e.

$$\int_{\mathcal{Y}} f(y^\epsilon)\rho^\infty(y^\epsilon)dy^\epsilon = 0. \tag{5.5}$$

Notice that for this particular model, the homogenized SDE (5.2) does not contain a drift coefficient. Indeed, from Eq.(2.11) the drift of the homogenized SDE is given by

$$F(X) \quad = \quad \int_{\mathcal{Y}} f_1(x^\epsilon, y^\epsilon)\rho^\infty(y^\epsilon; x^\epsilon)dy^\epsilon = 0$$

since $f_1(x^\epsilon, y^\epsilon) = 0$ in our case.

Similar to Chapter 4, the objective of this chapter is to show that the (ExtQV) for data from models of the form (5.1) is asymptotically unbiased for the diffusion coefficient of the limiting diffusion process. This result is summarized in the Theorem 5.3.

Before we proceed to the statement of our main result we should define x_n^ϵ the approximation of x^ϵ on a partition $\pi_n([0,T]) := \{t_0, \dots t_n\}$, $t_i := i\delta$.

Theorem 5.3. *Let $x^\epsilon(t) : [0,T] \to \mathbb{R}$ be a real–valued path described by Eq.(5.1). Then, subject to technical assumptions on the behavior of the functions f, g and β (given in Assumption 5.5 on page 50), as $\epsilon \to 0$ the square of the (ExtQV) is asymptotically unbiased for σ^2 in Eq.(5.2), i.e.*

$$\lim_{\epsilon\to 0}\lim_{n\to\infty} \mathbb{E}\left[\left(D_2^{Ext}(x_n^\epsilon)_T\right)^2\right] = 2\mathbb{E}\left[f(y^\epsilon)\Phi(y^\epsilon)\right]. \tag{5.6}$$

Steps of the proof.

1. As in the proof of the simple model in Chapter 4 and since the model of interest is again of bounded variation the expectation of the (ExtQV) can be obtained by computing the following expression

$$4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\Delta x_n^\epsilon(t_1)\Delta x_n^\epsilon(t_k)\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{1},\mathbf{k}}^{\mathbf{x}^\epsilon}\right)\right], \tag{5.7}$$

where $\mathcal{C}$, $c_{1,k}^{x^\epsilon}$ and $\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{1},\mathbf{k}}^{\mathbf{x}^\epsilon}\right)$ as before (see p.22).

2. Express the increments of the process x^ϵ in terms of $f(y^\epsilon(t_0))$ and $\Phi(y^\epsilon(t_0))$ and show that

$$\left|4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\Delta x_n^\epsilon(t_1)\Delta x_n^\epsilon(t_k)\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{1},\mathbf{k}}^{\mathbf{x}^\epsilon}\right)\right] - 2\mathbb{E}\left[f(y^\epsilon(t_0))\Phi(y^\epsilon(t_0))\right]\right| \to 0, \tag{5.8}$$

as $n \to \infty$ and $\epsilon \to 0$.

5.2 Analytical Proof of Theorem 5.3

In Lemma C.2 we have shown that by applying Itô–Taylor expansion on $f(y)$ (see Kloeden and Platen (1999, Chapter 5)) we get the following approximation for f

$$f\left(y_n^\epsilon(t)\right) = e^{\frac{(t-t_{i-1})}{\epsilon^2}\mathcal{L}_0} f\left(y_n^\epsilon(t_{i-1})\right) + \frac{1}{\epsilon}\int_{t_{i-1}}^{t} e^{\frac{(t-u)}{\epsilon^2}\mathcal{L}_0}(\nabla_{y_n^\epsilon} f\beta)(y_n^\epsilon(u))dV(u). \tag{5.9}$$

Given Eq.(5.9) we obtain the following expression for the increments of the process x_n^ϵ

$$\begin{aligned}\Delta x_n^\epsilon(t_i) &= \frac{1}{\epsilon}\int_{t_{i-1}}^{t_i} f\left(y_n^\epsilon(t)\right) dt \\ &= \epsilon\left(e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} - 1\right)\mathcal{L}_0^{-1} f(y_n^\epsilon(t_{i-1})) \\ &\qquad + \frac{1}{\epsilon^2}\int_{t_{i-1}}^{t_i}\int_{t_{i-1}}^{t} e^{\frac{(t-u)}{\epsilon^2}\mathcal{L}_0}(\nabla_{y_n^\epsilon} f\beta)(y_n^\epsilon(u))dV(u)dt. \quad (5.10)\end{aligned}$$

The justification of Eq.(5.10) can be found in Lemma C.3.

At this point, we have extracted an expression for the increments of the process x_n^ϵ similar to the one in Eq.(4.11) for the Ornstein–Uhlenbeck case. To simplify our computations, lets define

$$\psi\left(y_n^\epsilon(t_i)\right) := \epsilon\left(e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} - 1\right)\mathcal{L}_0^{-1} f(y_n^\epsilon(t_{i-1})), \quad (5.11)$$

and

$$M_{t_i} := \frac{1}{\epsilon^2}\int_{t_{i-1}}^{t_i}\int_{t_{i-1}}^{t} e^{\frac{(t-u)}{\epsilon^2}\mathcal{L}_0}(\nabla_{y_n^\epsilon} f\beta)(y(u))dV(u)dt. \quad (5.12)$$

Therefore, given the notation in equations (5.11) and (5.12), the product of our interest, $\Delta x_n^\epsilon(t_1)\Delta x_n^\epsilon(t_k)$, takes the following form

$$\begin{aligned}\Delta x_n^\epsilon(t_1)\Delta x_n^\epsilon(t_k) &= \psi\left(y_n^\epsilon(t_0)\right)\psi\left(y_n^\epsilon(t_{k-1})\right) + M_{t_1}M_{t_k} \\ &\qquad + \psi\left(y_n^\epsilon(t_{k-1})\right)M_{t_1} + \psi\left(y_n^\epsilon(t_0)\right)M_{t_k}, \quad (5.13)\end{aligned}$$

and the expectation in Eq.(5.7) becomes

$$\begin{aligned}\mathbb{E}\left[\Delta x_n^\epsilon(t_1)\Delta x_n^\epsilon(t_k)\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{1,k}}^{\mathbf{x_n^\epsilon}}\right)\right] &= \mathbb{E}\left[\psi\left(y_n^\epsilon(t_0)\right)\psi\left(y_n^\epsilon(t_{k-1})\right)\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi(y_n^\epsilon)} + \mathbf{M}\right)\right] \quad (5.14a) \\ &\quad + \mathbb{E}\left[M_{t_1}M_{t_k}\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi(y_n^\epsilon)} + \mathbf{M}\right)\right] \quad (5.14b) \\ &\quad + \mathbb{E}\left[\psi\left(y_n^\epsilon(t_{k-1})\right)M_{t_1}\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi(y_n^\epsilon)} + \mathbf{M}\right)\right] \quad (5.14c) \\ &\quad + \mathbb{E}\left[\psi\left(y_n^\epsilon(t_0)\right)M_{t_k}\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi(y_n^\epsilon)} + \mathbf{M}\right)\right], \quad (5.14d)\end{aligned}$$

where $\boldsymbol{\psi(y_n^\epsilon)} = \{\psi(y_n^\epsilon(t_0)), \ldots \psi(y_n^\epsilon(t_{k-1}))\}$ and $\mathbf{M} = \{M_{t_0}, \ldots, M_{t_{k-1}}\}$.

Our aim is to prove that

$$\underbrace{\left|4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\Delta x_n^\epsilon(t_1)\Delta x_n^\epsilon(t_k)\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{1,k}}^{\mathbf{x}_\mathbf{n}^\epsilon}\right)\right]-2\mathbb{E}\left[f\left(y_n^\epsilon(t_0)\right)\Phi\left(y_n^\epsilon(t_0)\right)\right]\right|}_{:=E} \quad\to\quad 0, \tag{5.15}$$

as $n\to\infty$ and $\epsilon\to 0$, or equivalently from Eq.(5.14),

$$\left|4\sum_{k=2}^{n}(n+1-k)\left[(5.14\text{a})+(5.14\text{b})+(5.14\text{c})+(5.14\text{d})\right]\right.$$
$$\left.-2\mathbb{E}\left[f\left(y_n^\epsilon(t_0)\right)\Phi\left(y_n^\epsilon(t_0)\right)\right]\right|\to 0. \tag{5.16}$$

But, from triangular inequality we get

$$\begin{aligned} E \quad\leq\quad & \left|4\sum_{k=2}^{n}(n+1-k)\left[(5.14\text{a})\right]-2\mathbb{E}\left[f\left(y_n^\epsilon(t_0)\right)\Phi\left(y_n^\epsilon(t_0)\right)\right]\right| \\ & +\left|4\sum_{k=2}^{n}(n+1-k)\left[(5.14\text{b})\right]\right|+\left|4\sum_{k=2}^{n}(n+1-k)\left[(5.14\text{c})\right]\right| \\ & +\left|4\sum_{k=2}^{n}(n+1-k)\left[(5.14\text{d})\right]\right|. \end{aligned} \tag{5.17}$$

In what follows, we initially prove that the second, third and fourth term in Eq.(5.17) tend to zero as $n\to\infty$ and later that the first term tend to zero as $n\to\infty$ and $\epsilon\to 0$.

Starting from the second term, applying consequently Jensen's, Cauchy–Schwarz inequality and the fact that $\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi}\left(\boldsymbol{y_n^\epsilon}\right)+\mathbf{M}\right)\leq 1$ we get

$$\begin{aligned} \left|4\sum_{k=2}^{n}(n+1-k)\left[(5.14\text{b})\right]\right| \quad &=\quad \left|4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[M_{t_1}M_{t_k}\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi}\left(\boldsymbol{y_n^\epsilon}\right)+\mathbf{M}\right)\right]\right| \\ &\leq\quad 4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\left|M_{t_1}M_{t_k}\right|\right] \\ &\leq\quad 4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[M_{t_1}^2\right]^{1/2}\mathbb{E}\left[M_{t_k}^2\right]^{1/2} \\ &\leq\quad \frac{4C}{\epsilon^7}\sum_{k=2}^{n}\frac{(n+1-k)}{n^3}\to 0, \end{aligned}$$

as $n \to \infty$ where the latter inequality comes from the result in Lemma C.4.

For the third term, to further simplify our computations, notice that

$$\begin{aligned}
\psi(y^\epsilon) &= \epsilon e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0}\mathcal{L}_0^{-1} f(y^\epsilon) - \epsilon \mathcal{L}_0^{-1} f(y^\epsilon) \\
&= \epsilon \sum_{m=0}^{\infty} \left(\frac{\delta}{\epsilon^2}\right)^m \frac{\mathcal{L}_0^{(m)} \mathcal{L}_0^{-1} f(y^\epsilon)}{m!} - \epsilon \mathcal{L}_0^{-1} f(y^\epsilon) \\
&= \frac{\delta}{\epsilon} f(y^\epsilon) + \epsilon \sum_{m=2}^{\infty} \left(\frac{\delta}{\epsilon^2}\right)^m \frac{\mathcal{L}_0^{(m)} \mathcal{L}_0^{-1} f(y^\epsilon)}{m!} \\
&= \frac{\delta}{\epsilon} f(y^\epsilon) - \underbrace{\epsilon \sum_{m=2}^{\infty} \left(\frac{\delta}{\epsilon^2}\right)^m \frac{\mathcal{L}_0^{(m)} \Phi(y^\epsilon)}{m!}}_{:=\lambda(y)} . \qquad (5.18)
\end{aligned}$$

Using the latter expression for $\psi(y)$ and the same arguments with the previous term we obtain

$$\begin{aligned}
\left| 4\sum_{k=2}^{n} (n+1-k) \left[(5.14\text{c})\right] \right| &= \left| 4\sum_{k=2}^{n} (n+1-k) \mathbb{E}\left[\left(\frac{\delta}{\epsilon} f(y_n^\epsilon(t_0)) - \lambda(y_n^\epsilon(t_0)) \right) \right.\right. \\
&\qquad \left.\left. \times M_{t_k} \mathbf{1}_{\mathcal{C}} \left(\boldsymbol{\psi}\left(\boldsymbol{y_n^\epsilon}\right) + \mathbf{M} \right) \right] \right| \\
&\leq \left| \frac{4}{\epsilon} \sum_{k=2}^{n} \frac{(n+1-k)}{n} \mathbb{E}\left[f(y_n^\epsilon(t_0)) M_{t_k} \times \right.\right. \\
&\qquad \left.\left. \mathbf{1}_{\mathcal{C}} \left(\boldsymbol{\psi}\left(\boldsymbol{y_n^\epsilon}\right) + \mathbf{M} \right) \right] \right| \qquad (5.19\text{a}) \\
&\quad + \left| 4\sum_{k=2}^{n} (n+1-k) \mathbb{E}\left[\lambda(y_n^\epsilon(t_0)) M_{t_k} \times \right.\right. \\
&\qquad \left.\left. \mathbf{1}_{\mathcal{C}} \left(\boldsymbol{\psi}\left(\boldsymbol{y_n^\epsilon}\right) + \mathbf{M} \right) \right] \right| . \qquad (5.19\text{b})
\end{aligned}$$

We treat each of the terms in Eq.(5.19a) and Eq.(5.19b) separately. For Eq.(5.19a), following the same approach as before we get

$$\text{Eq.(5.19a)} \quad \leq \quad C_\epsilon \sum_{k=2}^{n} \frac{(n+1-k)}{n} \mathbb{E}\left[f(y_n^\epsilon(t_0))^2 \right]^{1/2} \mathbb{E}\left[M_{t_k}^2 \right]^{1/2} .$$

where C_ϵ is a constant depending only on ϵ. Our assumptions on f and the result in Lemma C.4 suggest that the term in Eq.(5.19a) tends to zero as $n \to \infty$.

Similarly for Eq.(5.19b) we get

$$\text{Eq.(5.19b)} \quad \leq \quad C_\epsilon \sum_{k=2}^{n} (n+1-k)\mathbb{E}\left[\lambda(y_n^\epsilon(t_0))^2\right]^{1/2} \mathbb{E}\left[M_{t_k}^2\right]^{1/2},$$

and from Lemmas C.4 and C.5 we get

$$\text{Eq.(5.19b)} \quad \leq \quad C_\epsilon \sum_{k=2}^{n} \frac{(n+1-k)}{n^2 n^{3/2}} \to 0,$$

as $n \to \infty$.

Exactly the same procedure can be applied to show that the fourth term, Eq.(5.14d), also tends to zero as $n \to \infty$.

For the first term, E_{T_1}, one has

$$\begin{aligned} E_{T_1} \quad = \quad & \left| 4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\psi(y_n^\epsilon(t_0))\psi(y_n^\epsilon(t_0))\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi(y_n^\epsilon)} + \mathbf{M}\right)\right] \right. \\ & \left. -2\mathbb{E}\left[f\left(y_n^\epsilon(t_0)\right)\Phi\left(y_n^\epsilon(t_0)\right)\right] \right|. \end{aligned}$$

From triangular inequality

$$\begin{aligned} E_{T_1} \quad \leq \quad & \left| 4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\psi(y_n^\epsilon(t_0))\psi(y_n^\epsilon(t_0))\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi(y_n^\epsilon)}\right)\right] \right. \\ & \left. -2\mathbb{E}\left[f\left(y_n^\epsilon(t_0)\right)\Phi\left(y_n^\epsilon(t_0)\right)\right] \right| \qquad (5.20a) \\ & + \left| 4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\psi(y_n^\epsilon(t_0))\psi(y_n^\epsilon(t_{k-1}))\mathbf{1}_{\mathcal{C}}\left(|\mathbf{M}| - \left|\boldsymbol{\psi(y_n^\epsilon)}\right|\right)\right] \right|. \qquad (5.20b) \end{aligned}$$

Similar to previous steps and as $\mathbf{1}_{\mathcal{C}}\left(|\mathbf{M}| - \left|\boldsymbol{\psi(y_n^\epsilon)}\right|\right)$ implies that $\psi(y_n^\epsilon(t_{i-1})) \leq M_{t_i}$ $\forall i \in 1, \ldots, k$,

$$\begin{aligned} \text{Eq.(5.20b)} \quad & \leq C_\epsilon \sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[M_{t_1}^2\right]^{1/2}\mathbb{E}\left[M_{t_{k-1}}^2\right]^{1/2} \\ & \leq C_\epsilon \sum_{k=2}^{n} \frac{(n+1-k)}{n^3} \to 0 \end{aligned}$$

as $n \to \infty$.

To complete the proof of Theorem 5.3 it remains to show that Eq.(5.20a) tends to zero as $n \to \infty$ and $\epsilon \to 0$. The situation for this term is a bit more complicated than it was for the rest of the terms.

To begin with, using the expression for $\psi(y)$ given by Eq.(5.18), the Eq.(5.20a) takes the following form

$$
\begin{aligned}
\text{Eq.(5.20a)} = \Bigg| & 4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\psi(y_n^\epsilon(t_0))\psi(y_n^\epsilon(t_0))\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi}(\boldsymbol{y_n^\epsilon})\right)\right] \\
& \qquad - 2\mathbb{E}\left[f\left(y_n^\epsilon(t_0)\right)\Phi\left(y_n^\epsilon(t_0)\right)\right]\Bigg| \\
\leq \Bigg| & \frac{4}{\epsilon^2}\sum_{k=2}^{n}\frac{(n+1-k)}{n^2}\mathbb{E}\left[f(y_n^\epsilon(t_0))f(y_n^\epsilon(t_{k-1}))\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi}(\boldsymbol{y_n^\epsilon})\right)\right] \\
& + 4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\lambda(y_n^\epsilon(t_0))\lambda(y_n^\epsilon(t_{k-1}))\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi}(\boldsymbol{y_n^\epsilon})\right)\right] \\
& - \frac{4}{\epsilon}\sum_{k=2}^{n}\frac{(n+1-k)}{n}\mathbb{E}\left[f(y_n^\epsilon(t_0))\lambda(y_n^\epsilon(t_{k-1}))\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi}(\boldsymbol{y_n^\epsilon})\right)\right] \\
& - \frac{4}{\epsilon}\sum_{k=2}^{n}\frac{(n+1-k)}{n}\mathbb{E}\left[\lambda(y_n^\epsilon(t_0))f(y_n^\epsilon(t_{k-1}))\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi}(\boldsymbol{y_n^\epsilon})\right)\right] \\
& - 2\mathbb{E}\left[f\left(y_n^\epsilon(t_0)\right)\Phi\left(y_n^\epsilon(t_0)\right)\right]\Bigg|,
\end{aligned}
$$

and from triangular inequality,

$$
\begin{aligned}
\text{Eq.(5.20a)} \leq & \Bigg|4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[f(y_n^\epsilon(t_0))f(y_n^\epsilon(t_{k-1}))\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi}(\boldsymbol{y_n^\epsilon})\right)\right] \\
& \qquad - 2\mathbb{E}\left[f\left(y_n^\epsilon(t_0)\right)\Phi\left(y_n^\epsilon(t_0)\right)\right]\Bigg| && (5.21a) \\
& + \left|4\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\lambda(y_n^\epsilon(t_0))\lambda(y_n^\epsilon(t_{k-1}))\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi}(\boldsymbol{y_n^\epsilon})\right)\right]\right| && (5.21b) \\
& + \left|\frac{4}{\epsilon}\sum_{k=2}^{n}\frac{(n+1-k)}{n}\mathbb{E}\left[f(y_n^\epsilon(t_0))\lambda(y_n^\epsilon(t_{k-1}))\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi}(\boldsymbol{y_n^\epsilon})\right)\right]\right| && (5.21c) \\
& + \left|\frac{4}{\epsilon}\sum_{k=2}^{n}\frac{(n+1-k)}{n}\mathbb{E}\left[\lambda(y_n^\epsilon(t_0))f(y_n^\epsilon(t_{k-1}))\mathbf{1}_{\mathcal{C}}\left(\boldsymbol{\psi}(\boldsymbol{y_n^\epsilon})\right)\right]\right|. && (5.21d)
\end{aligned}
$$

Following the same procedure as for Eq.(5.17), we show that the terms (5.21b)–(5.21d) tend to zero as $n \to \infty$ and that the term (5.21a) tend to zero as $n \to \infty$ and $\epsilon \to 0$.

For the term (5.21b), the same approach as before and Lemma C.5 suggest that

$$\text{Eq.}(5.21\text{b}) \quad \leq \quad C_\epsilon \sum_{k=2}^{n} \frac{(n+1-k)}{n^4} \to 0,$$

as $n \to \infty$. Similarly, for the terms (5.21c) and (5.21d) we get

$$\text{Eq.}(5.21\text{c}) \quad \leq \quad C_\epsilon \sum_{k=2}^{n} \frac{(n+1-k)}{n^3} \to 0, \qquad \text{Eq.}(5.21\text{d}) \leq C_\epsilon \sum_{k=2}^{n} \frac{(n+1-k)}{n^3} \to 0,$$

as $n \to \infty$.

It remains to show that Eq.(5.21a) $\to 0$ as $n \to \infty$ and $\epsilon \to 0$.

In a similar way to Eq.(5.19) and by using the fact that $\psi(y) = \dfrac{\delta}{\epsilon} f(y) - \lambda(y)$, we obtain

$$\begin{aligned}
\text{Eq.}(5.21\text{a}) \quad = \quad & \Bigg| \frac{4}{\epsilon^2} \sum_{k=2}^{n} \frac{(n+1-k)}{n^2} \mathbb{E}\left[f(y_n^\epsilon(t_0)) f(y_n^\epsilon(t_{k-1})) \mathbf{1}_{\mathcal{C}} \left(\frac{\delta}{\epsilon} \mathbf{f}(\mathbf{y}_\mathbf{n}^\epsilon) - \boldsymbol{\lambda}(^\epsilon_{\boldsymbol{n}}) \right) \right] \\
& \qquad -2\mathbb{E}\left[f(y_n^\epsilon(t_0)) \Phi(y_n^\epsilon(t_0)) \right] \Bigg| \\
\leq \quad & \Bigg| \frac{4}{\epsilon^2} \sum_{k=2}^{n} \frac{(n+1-k)}{n^2} \mathbb{E}\left[f(y_n^\epsilon(t_0)) f(y_n^\epsilon(t_{k-1})) \mathbf{1}_{\mathcal{C}} \left(\mathbf{f}(\mathbf{y}_\mathbf{n}^\epsilon) \right) \right] \\
& \qquad -2\mathbb{E}\left[f(y_n^\epsilon(t_0)) \Phi(y_n^\epsilon(t_0)) \right] \Bigg| \\
& + \Bigg| \frac{4}{\epsilon^2} \sum_{k=2}^{n} \frac{(n+1-k)}{n^2} \mathbb{E}\left[f(y_n^\epsilon(t_0)) f(y_n^\epsilon(t_{k-1})) \mathbf{1}_{\mathcal{C}} \left(\frac{\epsilon}{\delta} \left| \boldsymbol{\lambda}(\boldsymbol{y}_\boldsymbol{n}^\epsilon) \right| - \left| \mathbf{f}(\mathbf{y}_\mathbf{n}^\epsilon) \right| \right) \right] \Bigg| \\
\leq \quad & \Bigg| \frac{4}{\epsilon^2} \sum_{k=2}^{n} \frac{(n+1-k)}{n^2} \mathbb{E}\left[f(y_n^\epsilon(t_0)) f(y_n^\epsilon(t_{k-1})) \mathbf{1}_{\mathcal{C}} \left(\mathbf{f}(\mathbf{y}_\mathbf{n}^\epsilon) \right) \right] \\
& \qquad -2\mathbb{E}\left[f(y_n^\epsilon(t_0)) \Phi(y_n^\epsilon(t_0)) \right] \Bigg| \\
& \qquad +4 \sum_{k=2}^{n} (n+1-k) \mathbb{E}\left[\lambda(y_n^\epsilon(t_0))^2 \right]^{1/2} \mathbb{E}\left[\lambda(y_n^\epsilon(t_{k-1}))^2 \right]^{1/2},
\end{aligned}$$

where the second term tends to zero as $n \to \infty$ (see Lemma (C.5)). Finally, to complete the proof of Theorem 5.3 we need to prove that the following result holds.

Proposition 5.4.

$$E_{T_1}^{n,\epsilon} : \quad = \quad \Bigg| \frac{4}{\epsilon^2} \sum_{k=2}^{n} \frac{(n+1-k)}{n^2} \mathbb{E}\left[f(y_n^\epsilon(t_0)) f(y_n^\epsilon(t_{k-1})) \mathbf{1}_{\mathcal{C}} \left(\mathbf{f}(\mathbf{y}_\mathbf{n}^\epsilon) \right) \right]$$

$$-2\mathbb{E}\left[f(y_n^\epsilon(t_0))\Phi(y_n^\epsilon(t_0))\right]\Bigg| \to 0. \tag{5.22}$$

as $n \to \infty$ and $\epsilon \to 0$.

Proof. Firstly, we treat the inner expectation

$$IE_{T_1}^{n,\epsilon} := \mathbb{E}\left[f(y_n^\epsilon(t_0))f(y_n^\epsilon(t_{k-1}))\mathbf{1}_{\mathcal{C}}\left(\mathbf{f}(\mathbf{y}_\mathbf{n}^\epsilon)\right)\right]. \tag{5.23}$$

Applying Tower property on Eq.(5.23) we get

$$\begin{aligned} IE_{T_1}^{n,\epsilon} &= \mathbb{E}\left[\mathbb{E}\left[f(y_n^\epsilon(t_0))f(y_n^\epsilon(t_{k-1}))\mathbf{1}_{\mathcal{C}}\left(\mathbf{f}(\mathbf{y}_\mathbf{n}^\epsilon)\right)\Big| y_n^\epsilon(t_0),\dots,y_n^\epsilon(t_{k-2})\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[f(y_n^\epsilon(t_0))f(y_n^\epsilon(t_{k-1}))\prod_{i=0}^{k-1}\mathbf{1}_{\mathcal{C}}\left(f(y_n^\epsilon(t_i))\right)\Big| y_n^\epsilon(t_0),\dots,y_n^\epsilon(t_{k-2})\right]\right]. \end{aligned} \tag{5.24}$$

(TOWIK) property gives

$$\begin{aligned} IE_{T_1}^{n,\epsilon} = \mathbb{E}\Bigg[f(y_n^\epsilon(t_0))\prod_{i=0}^{k-1}\mathbf{1}_{\mathcal{C}}\left(f(y_n^\epsilon(t_i))\right) \times \\ \mathbb{E}\left[f(y_n^\epsilon(t_{k-1}))\mathbf{1}_{\mathcal{C}}\left(f(y_n^\epsilon(t_{k-1}))\right)\Big| y_n^\epsilon(t_0),\dots,y_n^\epsilon(t_{k-2})\right]\Bigg]. \end{aligned} \tag{5.25}$$

Also, from Markov property,

$$IE_{T_1}^{n,\epsilon} = \mathbb{E}\left[f(y_n^\epsilon(t_0))\prod_{i=0}^{k-2}\mathbf{1}_{\mathcal{C}}\left(f(y_n^\epsilon(t_i))\right)\mathbb{E}\left[f(y_n^\epsilon(t_{k-1}))\mathbf{1}_{\mathcal{C}}\left(f(y_n^\epsilon(t_{k-1}))\right)\Big| y_n^\epsilon(t_{k-2})\right]\right]. \tag{5.26}$$

Now, let $K_{t_i} := \dfrac{1}{\epsilon}\displaystyle\int_{t_{i-1}}^{t_i} e^{\frac{(t_i-u)}{\epsilon^2}\mathcal{L}_0}(\nabla_{y_n^\epsilon} f\beta)(y_n^\epsilon(u))dV(u),\ i \in \{1,\dots k\}$. Then, from Eq.(5.9), the expectation

$$IIE_{T_1}^{n,\epsilon} := \mathbb{E}\left[f\left(y_n^\epsilon(t_{k-1})\right)\mathbf{1}_{\mathcal{C}}\left(f\left(y_n^\epsilon(t_{k-1})\right)\right)\Big| f y_n^\epsilon(t_{k-1})\right]$$

takes the following form

$$IIE_{T_1}^{n,\epsilon} = e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f\left(y_n^\epsilon(t_{k-2})\right)\mathbb{E}\left[\mathbf{1}_{\mathcal{C}}\left(f\left(y_n^\epsilon(t_{k-1})\right)\right)\Big| y_n^\epsilon(t_{k-2})\right]$$

$$
\begin{aligned}
& \qquad + \mathbb{E}\left[K_{t_{k-1}} \mathbf{1}_{\mathcal{C}} \left(f\left(y_n^{\epsilon}(t_{k-1}) \right) \right) \middle| y_n^{\epsilon}(t_{k-2}) \right] \\
= \quad & e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f\left(y_n^{\epsilon}(t_{k-2}) \right) - e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f\left(y_n^{\epsilon}(t_{k-2}) \right) \Phi_K \left(-e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f\left(y_n^{\epsilon}(t_{k-2}) \right) \right) \\
& \qquad + \sigma_K^2 \phi_K \left(e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f\left(y_n^{\epsilon}(t_{k-2}) \right) \right),
\end{aligned}
$$

where Φ_K is the cdf of $\{K_{t_i}\}_{i=1}^k$. Using the result in Lemma C.6 we obtain the following inequality

$$
IIE_{T_1}^{n,\epsilon} \quad \leq \quad e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f\left(y_n^{\epsilon}(t_{k-2}) \right) + \sigma_K^2 \phi_K \left(-e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f\left(y_n^{\epsilon}(t_{k-2}) \right) \right). \tag{5.27}
$$

Given Eq.(5.27), the Eq.(5.26) becomes

$$
\begin{aligned}
IE_{T_1}^{n,\epsilon} \quad \leq \quad & \mathbb{E}\left[f(y_n^{\epsilon}(t_0)) \prod_{i=0}^{k-2} \mathbf{1}_{\mathcal{C}} \left(f(y_n^{\epsilon}(t_i)) \right) e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f\left(y_n^{\epsilon}(t_{k-2}) \right) \right] \\
& + \sigma_K^2 \mathbb{E}\left[f\left(y_n^{\epsilon}(t_0) \right) \prod_{i=0}^{k-2} \mathbf{1}_{\mathcal{C}} \left(f\left(y_n^{\epsilon}(t_i) \right) \right) \phi_K \left(-e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f\left(y_n^{\epsilon}(t_{k-2}) \right) \right) \right].
\end{aligned}
$$

Following the same approach one more time one has

$$
\begin{aligned}
IE_{T_1}^{n,\epsilon} \quad \leq \quad & \mathbb{E}\left[f(y_n^{\epsilon}(t_0)) \prod_{i=0}^{k-3} \mathbf{1}_{\mathcal{C}} \left(f(y_n^{\epsilon}(t_i)) \right) e^{\frac{2\delta}{\epsilon^2}\mathcal{L}_0} f\left(y_n^{\epsilon}(t_{k-3}) \right) \right] \\
& + \sigma_K^2 e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} \mathbb{E}\left[f\left(y_n^{\epsilon}(t_0) \right) \prod_{i=0}^{k-3} \mathbf{1}_{\mathcal{C}} \left(f\left(y_n^{\epsilon}(t_i) \right) \right) \phi_K \left(-e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f\left(y_n^{\epsilon}(t_{k-3}) \right) \right) \right] \\
& + \sigma_K^2 \mathbb{E}\left[f\left(y_n^{\epsilon}(t_0) \right) \prod_{i=0}^{k-2} \mathbf{1}_{\mathcal{C}} \left(f\left(y_n^{\epsilon}(t_i) \right) \right) \phi_K \left(-e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f\left(y_n^{\epsilon}(t_{k-2}) \right) \right) \right].
\end{aligned}
$$

Inductively we get

$$
\begin{aligned}
IE_{T_1}^{n,\epsilon} \quad \leq \quad & \mathbb{E}\left[f(y_n^{\epsilon}(t_0)) e^{\frac{(k-1)\delta}{\epsilon^2}\mathcal{L}_0} f\left(y_n^{\epsilon}(t_0) \right) \right] & \text{(5.28a)} \\
& + \sigma_K^2 \sum_{j=0}^{k-2} e^{\frac{j\delta}{\epsilon^2}\mathcal{L}_0} \mathbb{E}\left[f\left(y_n^{\epsilon}(t_0) \right) \phi_K \left(-e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f\left(y_n^{\epsilon}(t_{k-2-j}) \right) \right) \right. \\
& \qquad \left. \times \prod_{i=0}^{k-2-j} \mathbf{1}_{\mathcal{C}} \left(f\left(y_n^{\epsilon}(t_i) \right) \right) \right]. & \text{(5.28b)}
\end{aligned}
$$

The first term, Eq.(5.28a), leads to the desired result. Indeed, going back to Eq.(5.22)

$$\begin{aligned} E_{T_1^{n,\epsilon}} &\leq 4\lim_{n\to\infty}\sum_{k=2}^{n}\frac{(n+1-k)}{(n\epsilon)^2}\mathbb{E}\left[f\left(y_n^\epsilon(t_0)\right)e^{\frac{(k-1)\delta}{\epsilon^2}\mathcal{L}_0}f\left(y_n^\epsilon(t_0)\right)\mathbf{1}_{\mathcal{C}}\left(f\left(y_n^\epsilon(t_0)\right)\right)\right] \\ &\to 2\mathbb{E}\left[f\left(y_n^\epsilon(t_0)\right)\Phi\left(y_n^\epsilon(t_0)\right)\right], \end{aligned} \tag{5.29}$$

as $n\to\infty$ and as $\epsilon\to 0$ which is what is required in Lemma 5.4. For a detailed justification of Eq.(5.29) see Lemma C.9.

Finally, from equations (5.22) and (5.28b), in order to finish the proof of Proposition 5.4 and consequently the proof of Theorem 5.3 we need to show that

$$\lim_{n\to\infty}\sum_{k=2}^{n}\frac{(n+1-k)}{(n\epsilon)^2}\,(\text{Eq.}(5.28\text{b}))=0.$$

Following the same approach with that in Chapter 4, see Eq.(4.32b), we obtain

$$\begin{aligned} E_{T_{12}^{n,\epsilon}} &:= \left|\sum_{k=2}^{n}\frac{(n+1-k)}{(n\epsilon)^2}\,(\text{Eq.}(5.28\text{b}))\right| \\ &\leq C_\epsilon\sum_{k=2}^{n}\frac{(n+1-k)}{n^2}\sigma_K^2\left[\phi_K\left(e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0}f\left(y_{k-2-j}\right)\right)^2\right]^{1/2} \\ &\quad\times\rho^{k-2}\sum_{j=0}^{k-2}\frac{\mathbb{E}\left[\left(e^{\frac{j\delta}{\epsilon^2}\mathcal{L}_0}f\left(y_n^\epsilon(t_0)\right)\right)^2\right]^{1/2}}{\rho^j}, \end{aligned}$$

where C_ϵ is a constant depending only on ϵ and $\rho=\mathbb{P}(f\left(y_n^\epsilon(t_{t_1})\right)>0, f\left(y_n^\epsilon(t_{t_1})\right)>0)^{1/2}$. From the result in Lemma C.7 we obtain

$$\begin{aligned} E_{T_{12}^{n,\epsilon}} &:= \left|\sum_{k=2}^{n}\frac{(n+1-k)}{(n\epsilon)^2}\,(\text{Eq.}(5.28\text{b}))\right| \\ &\leq C_\epsilon\sum_{k=2}^{n}\frac{(n+1-k)}{n^2}\sigma_K^2\mathbb{E}\left[\phi_K\left(e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0}f\left(y_n^\epsilon(t_{k-2-j})\right)\right)^2\right]^{1/2} \\ &\quad\times\rho^{k-2}\sum_{j=0}^{k-2}\frac{e^{\frac{cj\delta}{\epsilon^2}}}{\rho^j} \\ &= C_\epsilon\sum_{k=2}^{n}\frac{(n+1-k)}{n^2}\sigma_K^2\mathbb{E}\left[\phi_K\left(e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0}f\left(y_n^\epsilon(t)\right)\right)^2\right]^{1/2} \end{aligned}$$

$$\times \rho^{k-2} \sum_{j=0}^{k-2} \left(\frac{a_c}{\rho} \right)^j,$$

where $a_c = e^{\frac{c\delta}{\epsilon^2}}$ and c a positive constant.

Finally, given the technical assumption below, the situation is similar to the one in Chapter 4 and by Lemma B.6 we obtain that $E_{T_{12}^{n,\epsilon}} \to 0$ as $n \to \infty$ which concludes the proof of Proposition 5.4 and consequently the proof of Theorem 5.3. $\square$

Assumption 5.5 (Technical Assumptions for Theorem 5.3)**.** For the proof of Theorem 5.3 we require the followings

1. $$\mathbb{P}(f\left(y_n^\epsilon(t_1)\right) > 0, f\left(y_n^\epsilon(t_0)\right) > 0) = 1 - \mathcal{O}\left(\sqrt{\frac{\delta}{\epsilon^2}}\right),$$

2. $$\mathbb{E}\left[\left(\phi_K\left(e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f\left(y_n^\epsilon(t)\right)\right)\right)^2\right]^{1/2} \leq \mathcal{O}\left(\frac{1}{\sqrt{\sigma_K}}\right).$$

Hence, we have shown that under the appropriate assumptions, taking the expectation of the extrema quadratic variation on data generated by any model of the form in Eq.(5.1) gives in the limit as $n \to \infty$ and as $\epsilon \to 0$ the homogenized coefficient. In the following section we show that this is also the case when the corresponding homogenized equation has a drift coefficient.

5.3 Homogenization with Drift

In this section, we examine if our proposed estimator is affected when the corresponding homogenized SDE has a drift term. To do this we assume a multiscale model of the following form

$$dx^\epsilon(t) = \frac{1}{\epsilon} f\left(y^\epsilon(t)\right) dt + f_1\left(x^\epsilon(t)\right) dt, \qquad x^\epsilon(0) = x_0, \tag{5.30a}$$

$$dy^\epsilon(t) = \frac{1}{\epsilon^2} g(y^\epsilon(t)) dt + \frac{\beta(y^\epsilon(t))}{\epsilon} dV(t), \qquad y^\epsilon(0) = y_0. \tag{5.30b}$$

Then, the corresponding homogenized SDE is of the form

$$dX(t) = F(X)dt + \Sigma dW(t), \qquad X(0) = x_0, \tag{5.31}$$

where

$$F(X) = \int_{\mathcal{Y}} f_1(x^\epsilon)\rho^\infty(y^\epsilon)dy = f_1(x^\epsilon),$$

and Σ as before (see Eq.(2.14)).

Applying Itô's formula at Φ we obtain

$$d\Phi(y^\epsilon(t)) = -\frac{1}{\epsilon}dx^\epsilon(t)dt + \frac{1}{\epsilon}\left(\nabla_{y^\epsilon}\Phi\beta\right)(y^\epsilon(t))dV(t), \tag{5.32}$$

which implies that the increment of the x_n process have the following form

$$\Delta x_n^\epsilon(t_i) = -\epsilon\Delta\Phi_n(y_n^\epsilon(t_i)) + M_{t_i}, \tag{5.33}$$

where

$$M_{t_i} = \int_{t_{i-1}}^{t_i} \left(\nabla_{y^\epsilon}\Phi\beta\right)(y^\epsilon(t))dV(t).$$

As it can be noticed, Eq.(5.33) has the same form of that in Eq.(5.10) which corresponds to the no drift case. Therefore, the presence of the f_1 term in our model should not affect our estimator. Indeed, this statement is verified numerically in the Example 6.4 in Chapter 6.

5.4 Summary

In this chapter, we examined the performance of our proposed estimator for multiscale models given in Eq.(5.1) and Eq.(5.30). These models are of bounded variation and converge to a homogenized SDE without and with drift respectively. We proved that for both models, under the appropriate assumptions, the (ExtQV) results to an asymptotically unbiased estimator for the diffusion coefficient of the homogenized SDE.

CHAPTER 6

Simulation Study

In this chapter we consider four examples of multiscale systems that exhibit bounded variation. We present numerical evidence supporting that the (ExtQV) can be used for the efficient estimation of the homogenized diffusion coefficient. The first example is an immediate extension of the model examined in Chapter 4 and allow us to proceed to further generalization of our methodology which is further supported by the second example. The third example is in a slightly different context but our simulation study suggests that the (ExtQV) can still be used. The last example is the same as the one considered in Chapter 4 with the addition of drift in the slow variables of the model which results to a homogenized equation with drift.

6.1 Examples

Example 6.1. Consider the following fast/slow system of SDEs

$$dx = \frac{\sigma}{\epsilon} y^3 dt, \qquad x(0) = x_0, \tag{6.1a}$$

$$dy = -\frac{y}{\epsilon^2} dt + \frac{\sqrt{2}}{\epsilon} dV, \qquad y(0) = y_0, \tag{6.1b}$$

where V is the standard Brownian motion and initial conditions x_0 and y_0. The invariant density, ρ^∞, of the fast process in Eq.(6.1b) is the standard normal. For the model in Eq.(6.1) the centering condition is satisfied, namely

$$\int_{-\infty}^{\infty} f_0(y)\rho^\infty(y)dy = \sigma \int_{-\infty}^{\infty} y^3 \rho^\infty(y)dy = 0.$$

Without loss of generality, we assume $\sigma = 1$. The corresponding solution to the Poisson problem (see Definition 2.2) is

$$\Phi(y) = \frac{1}{3}y^3 + 2y. \tag{6.2}$$

Consequently, applying the theory in Chapter 2 the corresponding homogenized SDE is given by

$$dX = \left(2 \cdot \mathbb{E}\left[f(y)\Phi(y)\right]\right)^{1/2} dW = \sqrt{22}dW,$$

where W is a standard Brownian motion and is independent of V. Indeed, Eq.(2.14) suggests that the diffusion coefficient of the homogenized SDE is given by

$$\begin{aligned} \Sigma^2 &= 2\int_{-\infty}^{\infty} y^3 \cdot \left(\frac{1}{3}y^3 + 2y\right) \rho^{\infty}(y)dy, \\ &= 2 \cdot 11 = 22. \end{aligned} \tag{6.3}$$

The drift coefficient is zero since $f_1(x, y) = 0$ for this particular example.

Our objective is to examine if by applying the (ExtQV) to data generated by Eq.(6.1) we can effectively estimate the diffusion coefficient of the homogenized SDE (Eq.(6.3)).

As in the previous models examined so far, the evaluation of the expectation of the (ExtQV) requires to find an expression for the product $\Delta x_n(t_1)\Delta x_n(t_k)$ with respect to the initial condition y_0. It is easy to check that this is also the case for the model we examine here. For this reason, we only seek for an expression with respect to y_0 and ignore the rest of the terms.

Towards this direction, first notice that by applying Itô's formula at Φ, the solution to the Poisson problem, allows us to express the increment process $\Delta x_n(t_i)$ with respect to $\Delta\Phi_n(y(t_i))$. Indeed,

$$\begin{aligned} d\Phi(y) &= \frac{1}{\epsilon^2}\left(\mathcal{L}_0\Phi(y)\right)dt + \frac{1}{\epsilon}\nabla_y\Phi(y)dV, \\ &\overset{\text{Eq.(2.7)}}{=} -\frac{1}{\epsilon^2}y^3 dt + \frac{1}{\epsilon}\left(2 + y^2\right)dV, \end{aligned} \tag{6.4}$$

where $\mathcal{L}_0$ is the generator of the y process (Eq.(6.1b)) which is the same as the fast process in Chapter 4. Notice that Eq.(6.4) using Eq.(6.1a) can be expressed as

$$d\Phi(y) = -\frac{1}{\epsilon}dx + \frac{1}{\epsilon}\nabla_y\Phi(y)dV, \tag{6.5}$$

and therefore

$$dx = -\epsilon d\Phi(y) + \nabla_y \Phi(y) dV.$$

The latter results to the desired expression for the increment process $\Delta x_n(t_i)$ with respect to $\Delta\Phi_n(y(t_i))$, that is

$$\Delta x_n(t_i) = -\epsilon \Delta\Phi_n(y(t_i)) + \int_{t_{i-1}}^{t_i} \nabla_y \Phi(y)(t) dV(t). \tag{6.6}$$

Now, the objective is to express Eq.(6.6) with respect to the initial condition y_0. To do this, we first express $\Delta\Phi_n(y)$ with respect to y_0. Notice that $\Phi(y) = \frac{1}{3}y^3 + 2y$ implies that $f(y) = y^3 = 3(\Phi(y) - 2y)$. For simplicity set $m = 3$ and then Eq.(6.4) takes the following form

$$d\Phi(y) = -\frac{m}{\epsilon^2}\left(\Phi(y) - 2y\right) dt + \frac{1}{\epsilon} dM, \tag{6.7}$$

or, in integral form

$$\Phi(y(t)) = \Phi(y(0)) - \frac{m}{\epsilon^2}\int_0^t \left[\Phi(y(u)) - 2y(u)\right] du + \frac{1}{\epsilon}\int_0^t dM(u), \tag{6.8}$$

where $dM(t) := \nabla_y \Phi(y(t)) dV(t)$.

To obtain an approximation of the solution to Eq.(6.7) we apply simultaneously Picard iterations (see Ladroue and Papavasiliou (2013)) on the system of SDEs given by Eq.(6.8) and Eq.(6.1b).

For simplicity, set $X_t^{(1)} = t$, $X_t^{(2)} = V(t)$, $X_t^{(3)} = M(t)$, let also $\lambda = -\frac{1}{\epsilon^2}$ and $\mu = \frac{1}{\epsilon}$. On the interval $[0, \delta]$ and with initial guess

$$\Phi_{0,\delta}^{(0)} = \Phi(y_0), \qquad y_{0,\delta}^{(0)} = y_0, \tag{6.9}$$

after the first iteration we get

$$\begin{aligned} \Phi_{0,\delta}^{(1)} &= \Phi(y_0) + \lambda m \int_0^\delta \left(\Phi_{0,\delta}^{(0)} - 2y_{0,\delta}^{(0)}\right) dX_s^{(1)} + \mu \int_0^\delta dX_s^{(3)} \\ &\overset{\text{Eq.(6.9)}}{=} \Phi(y_0) + \lambda m \Phi(y_0) X_{0,\delta}^{(1)} - 2(\lambda m) y_0 X_{0,\delta}^{(1)} + \mu X_{0,\delta}^{(3)}. \end{aligned} \tag{6.10}$$

As we mentioned earlier, the desired result comes from the initial condition. For this reason, we ignore from now on the terms $X^{(2)}$ and $X^{(3)}$. Given that, at the second iteration we obtain

$$\begin{aligned}
\Phi_{0,\delta}^{(2)} &= \Phi(y_0) + (\lambda m)\Phi(y_0)X_{0,\delta}^{(1)} + (\lambda m)^2\Phi(y_0)X_{0,\delta}^{(1,1)} \\
&\quad -2(\lambda m)y_0X_{0,\delta}^{(1)} - 2\lambda^2(m^2+m)y_0X_{0,\delta}^{(1,1)}, \\
Y_{0,\delta}^{(2)} &= y_0 + \lambda y_0 X_{0,\delta}^{(1)} + \lambda^2 y_0 X_{0,\delta}^{(1,1)}.
\end{aligned}$$

Similarly, at the third and fourth iteration,

$$\begin{aligned}
\Phi_{0,\delta}^{(3)} &= \Phi(y_0) + (\lambda m)\Phi(y_0)X_{0,\delta}^{(1)} + (\lambda m)^2\Phi(y_0)X_{0,\delta}^{(1,1)} + (\lambda m)^3\Phi(y_0)X_{0,\delta}^{(1,1,1)} \\
&\quad -2(\lambda m)y_0X_{0,\delta}^{(1)} - 2\lambda^2(m^2+m)y_0X_{0,\delta}^{(1,1)} - 2\lambda^3(m^3+m^2+m)y_0X_{0,\delta}^{(1,1,1)}, \\
Y_{0,\delta}^{(3)} &= y_0 + \lambda y_0 X_{0,\delta}^{(1)} + \lambda^2 y_0 X_{0,\delta}^{(1,1)} + \lambda^3 y_0 X_{0,\delta}^{(1,1,1)},
\end{aligned}$$

and

$$\begin{aligned}
\Phi_{0,\delta}^{(4)} &= \Phi(y_0) + (\lambda m)\Phi(y_0)X_{0,\delta}^{(1)} + (\lambda m)^2\Phi(y_0)X_{0,\delta}^{(1,1)} + (\lambda m)^3\Phi(y_0)X_{0,\delta}^{(1,1,1)} \\
&\quad +(\lambda m)^4\Phi(y_0)X_{0,\delta}^{(1,1,1,1)} - 2(\lambda m)y_0X_{0,\delta}^{(1)} - 2\lambda^2(m^2+m)y_0X_{0,\delta}^{(1,1)} \\
&\quad -2\lambda^3(m^3+m^2+m)y_0X_{0,\delta}^{(1,1,1)} - 2\lambda^4(m^4+m^3+m^2+m)y_0X_{0,\delta}^{(1,1,1,1)}, \\
Y_{0,\delta}^{(4)} &= y_0 + \lambda y_0 X_{0,\delta}^{(1)} + \lambda^2 y_0 X_{0,\delta}^{(1,1)} + \lambda^3 y_0 X_{0,\delta}^{(1,1,1)} + \lambda^4 y_0 X_{0,\delta}^{(1,1,1,1)},
\end{aligned}$$

respectively.

After r iterations, one has

$$\Phi_{0,\delta}(r) = \Phi(y_0)\sum_{k=0}^{r}(\lambda m)^k X_{0,\delta}^{(\overbrace{1,1,\dots,1}^{k})} - 2y_0\sum_{k=1}^{r}\lambda^k\left[\sum_{i=1}^{k}m^i\right]X_{0,\delta}^{(\overbrace{1,1,\dots,1}^{k})}. \quad (6.11)$$

Given that,

$$\sum_{i=1}^{k} m^i = \frac{m(m^k-1)}{m-1}, \quad (6.12)$$

the Eq.(6.11) becomes

$$\Phi_{0,\delta}^{(r)} = \Phi(y_0)\sum_{k=0}^{r}(\lambda m)^k X_{0,\delta}^{(\overbrace{1,1,\dots,1}^{k})} - 2\frac{m}{m-1}y_0\sum_{k=1}^{r}\lambda^k\left(m^k-1\right)X_{0,\delta}^{(\overbrace{1,1,\dots,1}^{k})}. \quad (6.13)$$

Using the fact that $X_{0,\delta}^{(\overbrace{1,1,...,1}^{k})} = \frac{\delta^k}{k!}$, we get that as $r \to \infty$

$$\begin{aligned}
\Phi(y_0)\sum_{k=0}^{r}(\lambda m)^k X_{0,\delta}^{(\overbrace{1,1,...,1}^{k})} &\to \Phi(y_0)\sum_{k=0}^{\infty}(\lambda m)^k \frac{\delta^k}{k!} \\
&= \Phi(y_0)e^{\lambda m\delta} \\
&= \Phi(y_0)e^{-m\delta/\epsilon^2},
\end{aligned}$$

and

$$\begin{aligned}
y_0\sum_{k=1}^{r}\lambda^k\left(m^k-1\right)X_{0,\delta}^{(\overbrace{1,1,...,1}^{k})} &\to y_0\sum_{k=0}^{\infty}\lambda^k\left(m^k-1\right)\frac{\delta^k}{k!} \\
&= y_0\left(e^{-m\delta/\epsilon^2}-e^{-\delta/\epsilon^2}\right).
\end{aligned}$$

Putting everything together, one has

$$\Phi(y_\delta) = \Phi(y_0)e^{-m\delta/\epsilon^2} - 2\frac{m}{m-1}\left(e^{-m\delta/\epsilon^2}-e^{-\delta/\epsilon^2}\right)y_0, \tag{6.14}$$

or in general, at any interval $[t_{k-1}, t_k]$, and by setting $a = e^{-\delta/\epsilon^2}$ and using the fact that in this case $2\frac{m}{m-1} = m$ we get

$$\Phi(y(t_k)) = a^m\Phi(y(t_{k-1})) - m(a^m - a)y(t_{k-1}). \tag{6.15}$$

Equivalently, for $\Delta\Phi_n(y(t_k)) := \Phi(y(t_k)) - \Phi(y_{t_{k-1}})$

$$\Delta\Phi_n(y(t_k)) = (a^m - 1)\,\Phi(y(t_{k-1})) - m(a^m - a)y(t_{k-1}). \tag{6.16}$$

Also, recall that from the simple example where $f(y) = y$

$$y(t_k) = ay(t_{k-1}). \tag{6.17}$$

Notice that in Eq.(6.16) and Eq.(6.17) the terms involving the M and V have been ignored. This is due to the fact that as it has been shown in Chapter 4 the desired effect resulted form the initial condition and that the rest of terms tend to zero as $n \to \infty$. Based on that we focus on the examination of the deterministic part only.

Following the same approach as before, we use induction to express $\Delta\Phi_n$ in terms of the initial condition. At the first step we have

$$\Delta\Phi_n(y(t_k)) \;=\; a^m\,(a^m-1)\,\Phi(y(t_{k-1})) - m(a^m-a)\,[(a^m-1)+a]\,y(t_{k-1}).$$

Similarly,

$$\begin{aligned}
\Delta\Phi_n(y(t_k)) &= a^{2m}\,(a^m-1)\,\Phi(y(t_{k-2})) \\
&\qquad -m(a^m-a)\,[a^m(a^m-1)+a(a^m-1)+a]\,y(t_{k-2}) \\
\Delta\Phi_n(y(t_k)) &= a^{3m}\,(a^m-1)\,\Phi(y(t_{k-3})) - m(a^m-a)\times \\
&\quad \left[a^{2m}(a^m-1)+a^{m+1}(a^m-1)+a^2(a^m-1)+a^3\right]y(t_{k-3}) \\
\Delta\Phi_n(y(t_k)) &= a^{4m}\,(a^m-1)\,\Phi(y(t_{k-4})) - m(a^m-a)\left[a^{3m}(a^m-1)\right. \\
&\qquad \left. a^{2m+1}(a^m-1)+a^{m+2}(a^m-1)+a^3(a^m-1)+a^4\right]y(t_{k-4}) \\
&\;\;\vdots \\
\Delta\Phi_n(y(t_k)) &= a^{(k-1)m}\,(a^m-1)\,\Phi(y(t_0)). \\
&\qquad -m\,\underbrace{(a^m-a)\left[a^{k-1}+(a^m-1)\sum_{l=1}^{k-1}a^{m(k-1-l)+(l-1)}\right]}_{J}\,y(t_0).
\end{aligned} \tag{6.18}$$

The sum in Eq(6.18) is equal to

$$\sum_{l=1}^{k} a^{m(k-1-l)+(l-1)} = \frac{a^{(k-1)m}-a^{k-1}}{(a^m-a)}, \tag{6.19}$$

so that

$$\begin{aligned}
J &= (a^m-a)\left[a^{k-1}+(a^m-1)\sum_{l=1}^{k-1}a^{m(k-1-l)+(l-1)}\right] \\
&= a^{k-1}(a^m-a)+(a^m-1)\,(a^m-a)\,\frac{a^{(k-1)m}-a^{k-1}}{(a^m-a)} \\
&= a^{k-1}(a^m-a)+a^{(k-1)m}(a^m-1)-a^{(k-1)}\,(a^m-1) \\
&= a^{(k-1)}(a^m-a-a^m+1)+a^{(k-1)m}(a^m-1) \\
&= -a^{(k-1)}(a-1)+a^{(k-1)m}(a^m-1).
\end{aligned}$$

The latter implies that

$$-mJ = m\left[a^{(k-1)}\,(a-1)-a^{(k-1)m}\,(a^m-1)\right], \tag{6.20}$$

and therefore Eq.(6.18) can be expressed as

$$\Delta\Phi_n(y(t_k)) = a^{(k-1)m}\left(a^m-1\right)\Phi(y(t_0)) - m\left[a^{(k-1)m}\left(a^m-1\right) - a^{k-1}\left(a-1\right)\right]y(t_0). \tag{6.21}$$

At this point, we have managed to express $\Delta\Phi_n$ with respect to the initial condition. Substituting Eq.(6.21) into Eq.(6.6) and ignoring the martingale part for the reason explained earlier we get that

$$\begin{aligned}
\Delta x_n(t_1) &= -\epsilon\left[(a^m-1)\Phi(y(t_0)) - m\left[(a^m-1)-(a-1)\right]y(t_0)\right] \\
&= -\epsilon\left[(a^m-1)\Phi(y_0) - m\left[(a^m-a)\right]y(t_0)\right] \\
&\overset{\text{Eq.(6.2)}}{=} -\epsilon\frac{(a^m-1)}{m}f(y(t_0)) - \epsilon\left[m(a-1)-(a^m-a)\right]y(t_0) \\
&\approx -\epsilon\frac{(a^m-1)}{m}f(y(t_0)) - \epsilon\mathcal{O}\left(\frac{1}{(n\epsilon^2)^2}\right)y(t_0),
\end{aligned}$$

and

$$\begin{aligned}
\Delta x_n(t_k) = &-\epsilon a^{(k-1)m}\left(a^m-1\right)\Phi(y(t_0)) \\
&+\epsilon m\left[a^{(k-1)m}\left(a^m-1\right)-a^{k-1}\left(a-1\right)\right]y(t_0).
\end{aligned}$$

Using these expressions for $\Delta x_n(t_1)$ and $\Delta x_n(t_k)$ the expectation in Eq.(3.5) for this example becomes

$$\begin{aligned}
E = &\overbrace{2\epsilon^2\lim_{n\to\infty}\frac{(a^m-1)^2}{m}\sum_{k=2}^{n}(n+1-k)a^{(k-1)m}\mathbb{E}\left[f(y(t_0))\Phi(y(t_0))\right]}^{E_1} \\
&\underbrace{-2\epsilon^2\lim_{n\to\infty}(a^m-1)\sum_{k=2}^{n}(n+1-k)\left[a^{(k-1)m}\left(a^m-1\right)-a^{k-1}\left(a-1\right)\right]\mathbb{E}\left[y(t_0)\Phi(y(t_0))\right]}_{E_2} \\
&\overset{m=3}{=} 2\left[1-\frac{\epsilon^2}{3}\left(1-e^{-\frac{3}{\epsilon^2}}\right)\right]\mathbb{E}\left[f(y(t_0))\Phi(y(t_0))\right] \\
&\qquad -2\epsilon^2\left(2+e^{-\frac{3}{\epsilon^2}}-3e^{-\frac{1}{\epsilon^2}}\right)\mathbb{E}\left[y(t_0)\Phi(y(t_0))\right] \\
&\rightarrow 2\mathbb{E}\left[f(y(t_0))\Phi(y(t_0))\right], \qquad \text{as } \epsilon\to 0,
\end{aligned}$$

as required.

In general, for any $\sigma\in\mathbb{R}^+$

$$\lim_{n\to\infty}\mathbb{E}\left(D_2^{\text{Ext}}(x_n)_T^2\right) = 2\sigma^2\Bigg\{\left[1-\frac{\epsilon^2}{3}\left(1-e^{-\frac{3}{\epsilon^2}}\right)\right]\mathbb{E}\left[f(y(t_0))\Phi(y(t_0))\right]$$

$$
\begin{aligned}
&-\epsilon^2\left(2+e^{-\frac{3}{\epsilon^2}}-3e^{-\frac{1}{\epsilon^2}}\right)\mathbb{E}\left[y(t_0)\Phi(y(t_0))\right] \\
\to\quad & 2\sigma^2\mathbb{E}\left[f(y(t_0))\Phi(y(t_0))\right], \qquad \text{as } \epsilon\to 0.
\end{aligned}
$$

Notice that for the model we are considering in this example

$$
\begin{aligned}
\mathbb{E}\left[f(y)\Phi(y)\right] &= \int_{\mathbb{R}} y^3\left(\frac{y^3}{3}+2y\right)\rho^\infty(y)dy = 11, \\
\mathbb{E}\left[y\Phi(y)\right] &= \int_{\mathbb{R}} y_0\left(\frac{y^3}{3}+2y\right)\rho^\infty(y)dy = 3.
\end{aligned}
$$

Numerical Results

Similar to the numerical study in Chapter 4, we are going to investigate how the behavior of the (ExtQV) changes according to the choice of n, ϵ and σ.

Table 6.1 shows the values of the expectation of the (ExtQV) and its corresponding L_2–error when it is applied to the model (6.1). For this table we fix the value of $\sigma = 0.1$ and we consider five values of $\epsilon = (0.20, 0.15, 0.10, 0.05, 0.01)$ and four values of $n = \left(10^4, 10^5, 10^6, 10^7\right)$. The corresponding diffusion coefficient of the homogenized SDE in this case is $\Sigma^2 = 0.01\cdot 22 = 0.22$. As it can be seen from Table 6.1, as the value of n increases and the value of ϵ decreases, the expectation of the (ExtQV) tends to the real value of the homogenized diffusion coefficient. Furthermore, for decreasing n and ϵ the L_2–error decreases as well and tends to zero.

	ϵ	0.20	0.15	0.10	0.05	0.01
$n = 10^4$	$\mathbb{E}\left[D_2^{\text{Ext}}(x_n)_T^2\right]$	0.1938	0.2087	0.2177	0.2368	1.6072
	L_2-error	0.1260	0.0486	0.0237	0.0071	1.9342
$n = 10^5$	$\mathbb{E}\left[D_2^{\text{Ext}}(x_n)_T^2\right]$	0.2059	0.2086	0.2176	0.2281	0.2638
	L_2-error	0.0860	0.0611	0.0257	0.0078	0.0023
$n = 10^6$	$\mathbb{E}\left[D_2^{\text{Ext}}(x_n)_T^2\right]$	0.1927	0.2047	0.2169	0.2217	0.2284
	L_2-error	0.0630	0.0439	0.0268	0.0071	0.0004
$n = 10^7$	$\mathbb{E}\left[D_2^{\text{Ext}}(x_n)_T^2\right]$	0.2119	0.2075	0.211	0.2330	0.2209
	L_2-error	0.0986	0.0434	0.0235	0.0073	0.0003

Table 6.1: Expectation and L_2–error of the (ExtQV) for different ϵ's, n's and for $\sigma = 0.10$.

Similar to our studies in Figures 4.2 and 4.3, here we fix again $n = 10^6$ and investigate

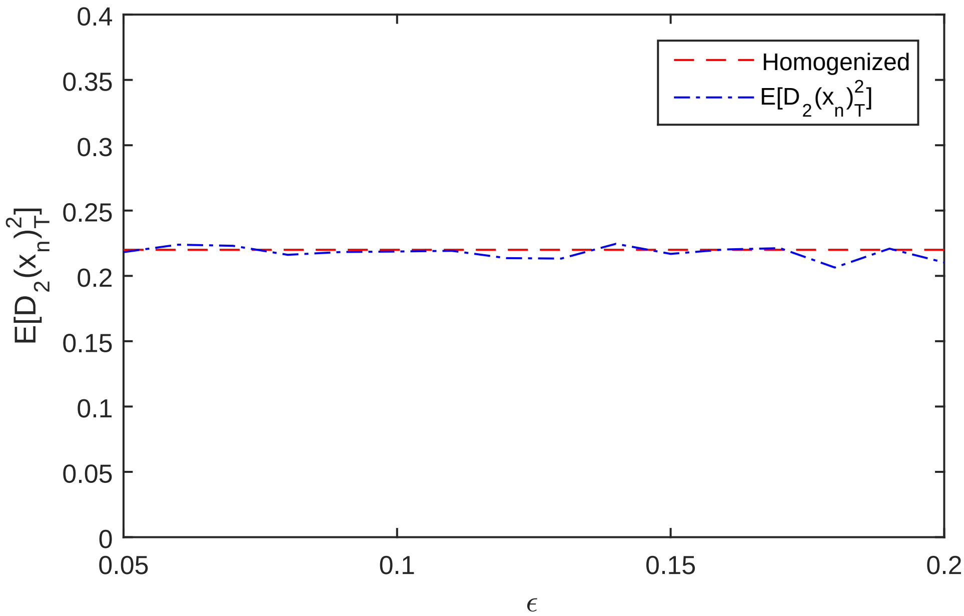

Figure 6.1: Expectation of the (ExtQV) for different ϵ's and for fixed $n = 10^6$ and $\sigma = 0.1$.

the behavior of the (ExtQV) with respect to the scale separation parameter ϵ for $\sigma = 0.10$, and $\sigma = 0.50$. The corresponding plots are given in Figures 6.1 and 6.2. For the former case, the expectation of the (ExtQV) is going to 0.22 for small values of ϵ and for the latter case is going to 5.50. Therefore as $\epsilon \to 0$ in both cases the expectation of the (ExtQV) approaches the actual value of the corresponding homogenized diffusion coefficient.

The approach followed so far can be easily extended to any multiscale system whose fast dynamics are driven by an Ornstein–Uhlenbeck process of the form in Eq.(5.1b) and with slow dynamics of the following form

$$dx(t) = \frac{f_m(y(t))}{\epsilon} dt, \tag{6.22}$$

where

$$f_m(y) = \begin{cases} y^m & \text{if } m \text{ is odd,} \\ 2^{-(m/2)} \dfrac{m!}{(m/2)!} - y^m & \text{if } m \text{ is even.} \end{cases} \tag{6.23}$$

The form of the function f_m as given in Eq.(6.23) reassures that the centering condition (Assumption 2.5) is satisfied. For this class of models the solution to the

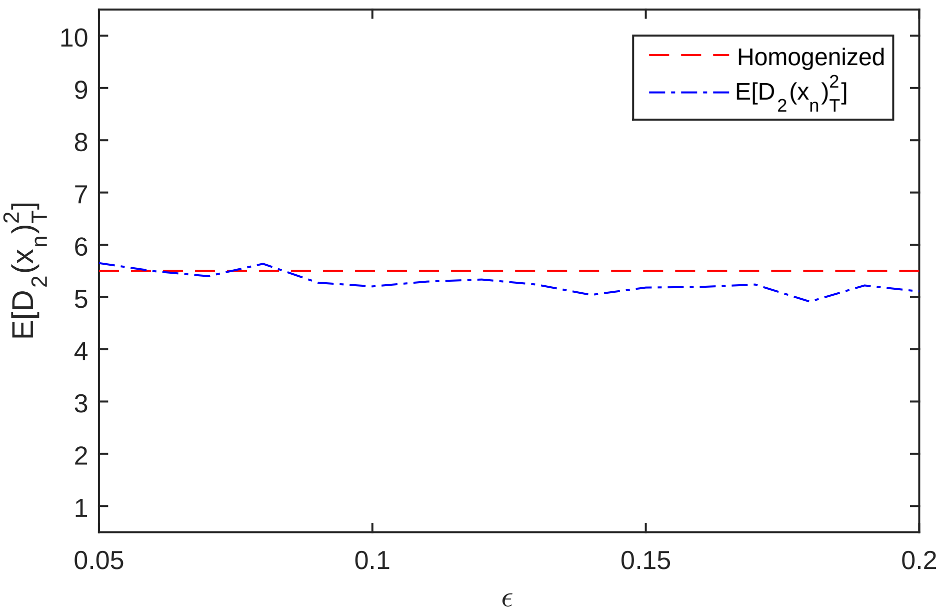

Figure 6.2: Expectation of the (ExtQV) for different ϵ's and for fixed $n = 10^6$ and $\sigma = 0.5$.

Poisson problem has the following form

$$\Phi_m(y) = \begin{cases} \dfrac{1}{m} f_m(y) + \displaystyle\sum_{j=1}^{\frac{m-1}{2}} \dfrac{\prod_{k=1}^{\frac{m+1}{2}-j} (m-(2k-1))}{(2j-1)} f_{(2j-1)}(y), & m \geq 3, \text{ and } m \text{ odd}, \\ \dfrac{1}{m} f_m(y) + \displaystyle\sum_{j=1}^{\frac{m}{2}-1} \dfrac{\prod_{k=1}^{\frac{m}{2}-j} (m-(2k-1))}{2j} f_{(2j)}(y), & m \geq 2, \text{ and } m \text{ even}. \end{cases} \tag{6.24}$$

Following the same approach as before, we can easily extend the results from the Example 6.1 to the class of problems with f being of the form in Eq.(6.23). In the example below, we illustrate numerically the above statement.

Example 6.2. Consider the following multiscale system of SDEs

$$\begin{aligned} dx &= \frac{\sigma}{\epsilon}\left(1-y^2\right) dt, \\ dy &= -\frac{1}{\epsilon^2} y dt + \frac{\sqrt{2}}{\epsilon} dV, \end{aligned}$$

where V is the standard Browian motion and initial conditions x_0 and y_0. The corresponding solution to the Poisson problem is $\Phi(y) = \dfrac{1}{2}(1-y^2)$ and consequently, applying the theory reviewed in Chapter 2, the homogenized SDE has the following form

$$dX = \sigma\sqrt{2}dW. \tag{6.25}$$

Following the same approach as in the previous section, one has

$$\begin{aligned} d\Phi(y) &= -\frac{1}{\epsilon^2}f(y)dt + \frac{1}{\epsilon}dM & (6.26)\\ &= -\frac{2}{\epsilon^2}\Phi(y)dt + \frac{1}{\epsilon}dM, & (6.27)\end{aligned}$$

where $dM = \sqrt{2}\nabla_y\Phi(y)dV$. Picard iteration suggests that an approximation to the solution of the SDE is given by

$$\Phi(y(t_k)) = e^{-\frac{2\delta}{\epsilon^2}}\Phi(y(t_{k-1})),$$

and the increments of Φ with respect to the initial condition can be expressed as

$$\Delta\Phi_n(y(t_k)) = a^{2(k-1)}(a^2-1)\Phi(y_0), \qquad \text{with } a = e^{-\frac{\delta}{\epsilon^2}}.$$

As before, here we also ignore the martingale part. Hence, from Eq.(6.26) the increments of the process x can be expressed by

$$\begin{aligned} \Delta x_n(t_1) &= -\epsilon(a^2-1)\Phi(y_0)\\ &= -\epsilon\frac{(a^2-1)}{2}f(y_0), & (6.28)\\ \Delta x_n(t_k) &= -\epsilon a^{2(k-1)}(a^2-1)\Phi(y_0). & (6.29)\end{aligned}$$

For this particular example, the expectation

$$E_n := 2\mathbb{E}\left[\sum_{k=2}^{n}(n+1-k)\Delta x_n(t_1)\Delta x_n(t_1)\mathbf{1}(C^{\mathbf{x}}_{i,j})\right],$$

given Eq.(6.28) and Eq.(6.29) becomes

$$\begin{aligned} E_n &= -2\epsilon^2\frac{(a^2-1)^2}{2}\sum_{k=2}^{n}(n+1-k)a^{2(k-1)}\mathbb{E}\left[f(y_0)\Phi(y_0)\right]\\ &= -\epsilon^2(a^2-1)^2\sum_{k=2}^{n}(n+1-k)a^{2(k-1)}\mathbb{E}\left[f(y_0)\Phi(y_0)\right].\end{aligned}$$

Then,

$$\lim_{n\to\infty} E_n = \left[2-\epsilon^2\left(1-e^{-\frac{2}{\epsilon^2}}\right)\right]\mathbb{E}\left[f(y_0)\Phi(y_0)\right]$$

$$
\begin{aligned}
&= \left[2-\epsilon^2\left(1-e^{-\frac{2}{\epsilon^2}}\right)\right]\int_{-\infty}^{\infty} f(y_0)\Phi(y_0)\rho^{\infty}(y_0)dy_0 \\
&= \left[2-\epsilon^2\left(1-e^{-\frac{2}{\epsilon^2}}\right)\right]\int_{-\infty}^{\infty} \frac{(1-y_0^2)^2}{2}\rho^{\infty}(y_0)dy_0 \\
&= 2-\epsilon^2\left(1-e^{-\frac{2}{\epsilon^2}}\right).
\end{aligned}
$$

Finally,

$$
\lim_{\epsilon\to 0}\lim_{n\to\infty} E_n = 2 = \Sigma^2, \tag{6.30}
$$

which is the diffusion coefficient of the limiting diffusion process for this example and for $\sigma = 1$ (see Eq.(6.25)).

Similarly to what we have done in previous examples, we examine the (ExtQV) for four values of n and five values of ϵ and the results are shown in Table 6.2 and Figure 6.3. First, as it was also noticed in the examples considered so far, we should consider a time step $\delta = T/n$ sufficient small compared to ϵ^2 in order the error due to discretization to become negligible. Indeed, for $n = 10^4$ and for $\epsilon = 0.10$ we can see that the performance of our estimator is not satisfactory. However, both Table

	ϵ	0.20	0.15	0.10	0.05	0.01
$n = 10^4$	$\mathbb{E}\left[D_2^{\text{Ext}}(x_n)_T^2\right]$	2.0426	2.0364	2.2174	2.3837	16.6118
	L_2-error	5.4799	3.0374	2.1654	0.5906	213.9965
$n = 10^5$	$\mathbb{E}\left[D_2^{\text{Ext}}(x_n)_T^2\right]$	1.9373	1.9752	2.0926	2.1041	2.6470
	L_2-error	4.2892	3.2262	1.4939	0.3571	0.4380
$n = 10^6$	$\mathbb{E}\left[D_2^{\text{Ext}}(x_n)_T^2\right]$	1.9120	2.0136	1.9397	2.0416	2.1744
	L_2-error	4.9281	2.9718	1.2007	0.3632	0.0463
$n = 10^7$	$\mathbb{E}\left[D_2^{\text{Ext}}(x_n)_T^2\right]$	1.9312	1.9721	2.0525	2.0061	2.0546
	L_2-error	6.3127	3.3509	1.6203	0.3273	0.0176

Table 6.2: Expectation and L_2–error of the (ExtQV) for different ϵ's, n's and for $\sigma = 1$.

6.2 and Figure 6.3 suggest, that by keeping ϵ fixed and increasing n the expectation of the (ExtQV) converges to 2 (the diffusion coefficient in this case). Also, keeping fixed $n = 10^7$ and decreasing ϵ we see that the L_2–error approaches the zero.

In the next example, we modify our context in the sense that the fast dynamics are not described by an (OU) process.

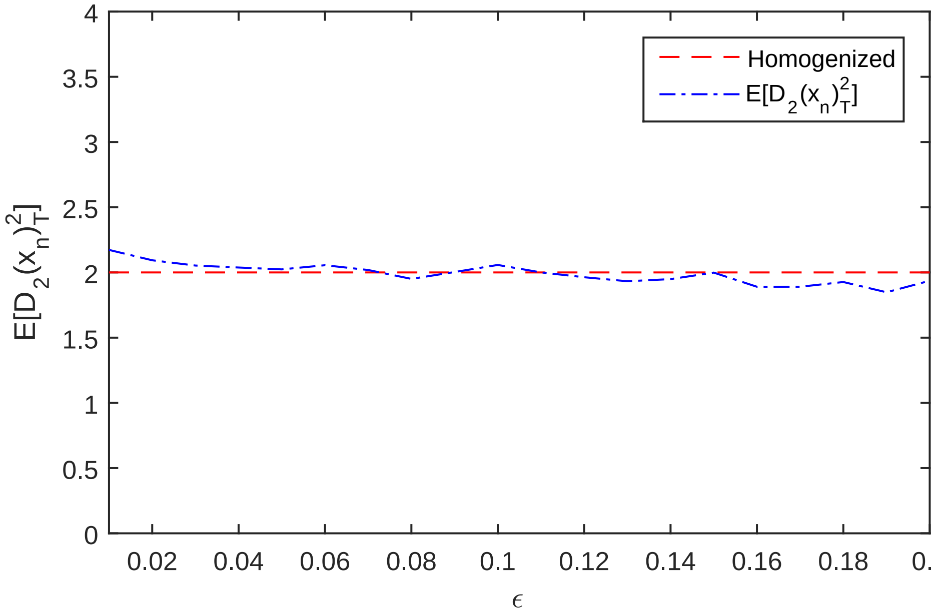

Figure 6.3: Expectation of the (ExtQV) for different ϵ's and fixed $n = 10^6$ and $\sigma = 1$.

Example 6.3. Consider the following multiscale system

$$dx = \sigma \frac{\sin(y)}{\epsilon} dt, \tag{6.31a}$$

$$dy = -\frac{\sin(y)}{\epsilon^2} dt + \frac{1}{\epsilon} dW. \tag{6.31b}$$

It is easy to see that the corresponding homogenized SDE is

$$dX = \sigma dW. \tag{6.32}$$

Indeed, from Eq.(6.31)

$$-\sigma \epsilon dy = \frac{\sigma}{\epsilon} \sin(y) dt - \sigma dW = \frac{1}{\epsilon} dx - \sigma dW$$

which implies that as $\epsilon \to 0$

$$dX = \sigma dW.$$

Table 6.3 illustrates the expectation of the (ExtQV) and its corresponding L_2–error when applied to the model (6.31) for $\sigma = \sqrt{0.5}$. As in the previous examples, we consider five values of $\epsilon = (0.20, 0.15, 0.10, 0.05, 0.01)$ and three values of $n =$

$(10^4, 10^5, 10^6)$. For $\sigma = \sqrt{0.5}$, the corresponding homogenized diffusion coefficient is equal to 0.5. Similarly to the previous examples, we observe that as the value of n increases and the value of ϵ decreases both the expectation of the (ExtQV) and the L_2–error tend to the desired quantity, that is the real value of the homogenized and coefficient and zero respectively.

	ϵ	0.20	0.15	0.10	0.05	0.01
$n = 10^4$	$\mathbb{E}\left[D_2^{\text{Ext}}(x_n)_T^2\right]$	0.3889	0.4023	0.4320	0.4604	0.8046
	L_2–error	0.0571	0.0391	0.0198	0.0059	0.0932
$n = 10^5$	$\mathbb{E}\left[D_2^{\text{Ext}}(x_n)_T^2\right]$	0.3707	0.3878	0.3974	0.4055	0.4336
	L_2–error	0.0599	0.0408	0.0243	0.0123	0.0046
$n = 10^6$	$\mathbb{E}\left[D_2^{\text{Ext}}(x_n)_T^2\right]$	0.3761	0.4005	0.4064	0.4201	0.4990
	L_2–error	0.0577	0.0393	0.0228	0.0102	0.0002

Table 6.3: Expectation and L_2–error of the (ExtQV) for different ϵ's, n's and $\sigma = \sqrt{0.5}$.

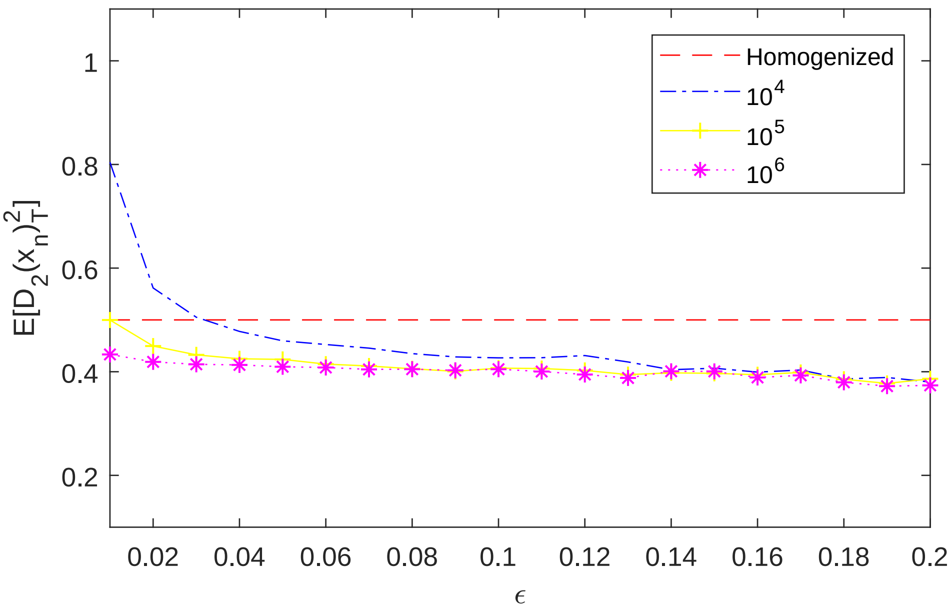

Figure 6.4: Expectation of the (ExtQV) for different ϵ's and n's and for $\sigma^2 = 0.5$.

The same conclusion can be extracted by observing Figures 6.4 and 6.5.

Finally, in the example below we demonstrate that our proposed estimator can be also applied in cases where the corresponding homogenized equation contains a drift

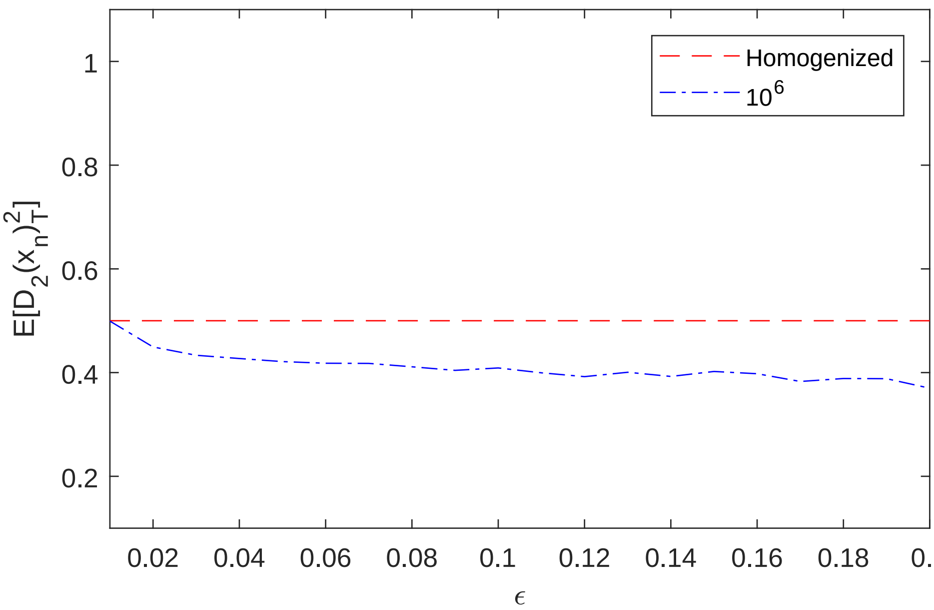

Figure 6.5: Expectation of the (ExtQV) for different ϵ's, constant $n = 10^6$ and for $\sigma = 0.5$.

term. In other words, we consider a multiscale system of the form in Eq.(5.30).

Example 6.4. Consider the following fast/slow system

$$dx = \frac{\sigma}{\epsilon} y dt + \sin(x) dt, \tag{6.33a}$$

$$dy = -\frac{1}{\epsilon^2} y dt + \frac{1}{\epsilon} dV. \tag{6.33b}$$

The corresponding homogenized SDE is

$$dX = \sin(X) dt + \sigma dW. \tag{6.34}$$

Similar numerical studies are performed for this model and Table 6.4 shows the Expectation and L_2–error of the Extrema Quadratic Variation for the same values of n and ϵ considered in the previous examples. Based on Table 6.4 and Figure 6.6 we conclude that the drift coefficient does not affect our proposed estimator.

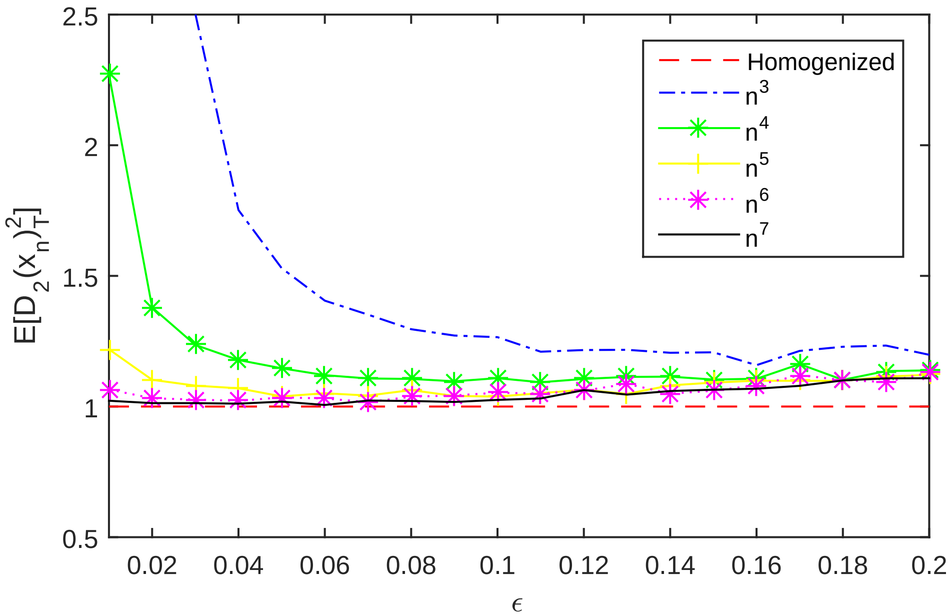

Figure 6.6: Expectation of the (ExtQV) for the model in Example 6.4 for different ϵ's, n's and for $\sigma = 1$.

	ϵ	0.20	0.15	0.10	0.05	0.01
$n = 10^4$	$\mathbb{E}\left[D_2^{\text{Ext}}(x_n)_T^2\right]$	1.1429	1.1112	1.1039	1.1289	2.2738
	L_2-error	0.4890	0.3055	0.1462	0.0505	1.6949
$n = 10^5$	$\mathbb{E}\left[D_2^{\text{Ext}}(x_n)_T^2\right]$	1.1258	1.0709	1.0559	1.0491	1.2194
	L_2-error	0.5250	0.2695	0.1312	0.0.12	0.0497
$n = 10^6$	$\mathbb{E}\left[D_2^{\text{Ext}}(x_n)_T^2\right]$	1.1488	1.0808	1.0282	1.0446	1.2196
	L_2-error	0.5822	0.2649	0.1115	0.0439	0.0164
$n = 10^7$	$\mathbb{E}\left[D_2^{\text{Ext}}(x_n)_T^2\right]$	1.1476	1.0728	1.0430	1.0595	1.2185
	L_2-error	0.5661	0.2778	0.1097	0.0356	0.0167

Table 6.4: Expectation and L_2–error of the (ExtQV) for the model in Example 6.4 for different ϵ's, n's and for $\sigma = 1$.

6.2 Summary

In this chapter, we considered four different examples of multiscale systems with zero quadratic variation. For all four of them, it was examined if the (ExtQV) can be applied to the corresponding multiscale data to effectively estimate the diffusion coefficient of the limiting diffusion process. Both theoretical and numerical evidence suggest that the (ExtQV) is asymptotically unbiased for the diffusion coefficient of the homogenized process, i.e., we can use the (ExtQV) to correctly identify the desired quantity in the limit of $n \to \infty$. Furthermore, it was illustrated numerically that our proposed estimator performs with an L_2–error of order ϵ^2 which is very satisfactory compared with other estimators in literature.

CHAPTER 7

Extrema Quadratic Variation for Multiscale Diffusions with Bounded Quadratic Variation

This chapter explores if our proposed estimator can also be applied to multiscale data that exhibit bounded non–zero quadratic variation. In order to do this, we consider the model introduced in Chapter 4 with an additional noise term to the slow variables of the model. The model of consideration now has the following form

$$dx^\epsilon(t) = \frac{1}{\epsilon} y^\epsilon(t)dt + dW(t), \qquad x^\epsilon(0) = x_0, \tag{7.1a}$$

$$dy(t) = -\frac{1}{\epsilon^2} y^\epsilon(t)dt + \frac{1}{\epsilon} dV(t), \qquad y^\epsilon(0) = y_0, \tag{7.1b}$$

where W and V are independent standard Brownian motions.

The additional Brownian motion term, W_t, in the slow variables of the model rises the question of how the (ExtQV) behaves when it is applied to Brownian motion paths. For this reason, in the following section we compute the expectation of the (ExtQV) for the Brownian motion.

7.1 Extrema Quadratic Variation for the Brownian Motion

The following result holds for the (ExtQV) of the Brownian motion process.

Proposition 7.1. *Let $W_n : [0,T] \to \mathbb{R}$ be a realization of a Brownian motion path on $[0,T]$. The (ExtQV) of the Brownian motion process is asymptotically equal to* *quantity* $T\left(1+\frac{4}{\pi}\right)$, *i.e.*

$$\lim_{n\to\infty}\mathbb{E}\left[D_2^{Ext}\left(W_n\right)_T^2\right]=T\left(1+\frac{4}{\pi}\right). \tag{7.2}$$

Proof. Using expression in Eq.(3.4) we obtain

$$\mathbb{E}\left[D_2^{\text{Ext}}\left(W_n\right)_T^2\right] \;=\; \mathbb{E}\left[D_2\left(W_n\right)_T^2\right]+2\sum_{i=2}^{n}\sum_{j=1}^{i-1}\mathbb{E}\left[\Delta W_n(t_i)\Delta W_n(t_j)\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{i,j}}^{\mathbf{W}}\right)\right].$$

where $\mathbf{c}_{\mathbf{i,j}}^{\mathbf{W}}$ is as defined in Chapter 3, i.e.

$$\mathbf{c}_{\mathbf{i,j}}^{\mathbf{W}}=\Delta W_n(t_i)\Delta_n W(t_j)>0,\ldots,\Delta W_n(t_i)\Delta_n W(t_{i-1})>0. \tag{7.3}$$

From the classical stochastic calculus theory (see Karatzas and Shreve (2012)), it is very well known that the quadratic variation of the Brownian motion equals to T in the limit as $n\to\infty$. Also, recall that the increments of the Brownian motion are independent Gaussian random variables with zero mean and variance equals to $\delta := T/n$.

For simplicity, we set $\eta_i := \dfrac{\Delta W_n(t_i)}{\sqrt{\delta}}$, then $\eta_i \overset{i.i.d}{\sim} \mathcal{N}(0,1)$ and $\Delta W_n(t_i)\Delta W_n(t_j) = \delta\eta_i\eta_j$. The limit of the expectation of the (ExtQV) for the Brownian motion is then

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}\left[D_2^{\text{Ext}}\left(W_n\right)_T^2\right] &= T+2\lim_{n\to\infty}\frac{1}{n}\sum_{i=2}^{n}\sum_{j=1}^{i-1}\mathbb{E}\left[\eta_i\eta_j\mathbf{1}_{\mathcal{C}}\left(\mathbf{c}_{\mathbf{i,j}}^{\mathbf{W}}\right)\right]\\ &= T+2\lim_{n\to\infty}\frac{1}{n}\sum_{i=2}^{n}\sum_{j=1}^{i-1}\mathbb{E}\left[\eta_i\eta_j\middle|\mathbf{c}_{\mathbf{i,j}}^{\mathbf{W}}\right]\mathbb{P}\left[\mathbf{c}_{\mathbf{i,j}}^{\mathbf{W}}\right],\end{aligned} \tag{7.4}$$

or equivalently, by stationarity,

$$\lim_{n\to\infty}\mathbb{E}\left[D_2^{\text{Ext}}\left(W_n\right)_T^2\right] \;=\; T+2\lim_{n\to\infty}\frac{1}{n}\sum_{k=2}^{n}(n+1-k)\mathbb{E}\left[\eta_1\eta_k\middle|\mathbf{c}_{\mathbf{k,1}}^{\mathbf{W}}\right]\mathbb{P}\left[\mathbf{c}_{\mathbf{k,1}}^{\mathbf{W}}\right]. \tag{7.5}$$

From Eq.(7.3), the probability in Eq.(7.5) is given by,

$$\mathbb{P}\left[\mathbf{c}_{\mathbf{k,1}}^{\mathbf{W}}\right] \;=\; \mathbb{P}\left[\eta_i\eta_j>0,\ldots\eta_i\eta_{i-1}>0\right]$$

$$= \quad \mathbb{P}\left[\eta_j > 0, \dots \eta_i > 0\right] + \mathbb{P}\left[\eta_j < 0, \dots \eta_i < 0\right]. \tag{7.6}$$

Due to symmetry for Gaussian random variables, the two probabilities at the (RHS) of Eq.(7.6) are equal. Hence,

$$\begin{aligned} \mathbb{P}\left[\mathbf{c}^{\mathbf{W}}_{\mathbf{k},\mathbf{1}}\right] &= 2\mathbb{P}\left[\eta_j > 0, \dots \eta_i > 0\right] \\ &= 2\mathbb{P}\left[\eta_j < 0, \dots \eta_i < 0\right], \end{aligned}$$

and

$$\begin{aligned} \lim_{n\to\infty} \mathbb{E}\left[D_2^{\text{Ext}}\left(W_n\right)_T^2\right] &= T + 4 \lim_{n\to\infty} \frac{1}{n} \sum_{i=2}^{n} \sum_{j=1}^{i-1} \mathbb{E}\left[\eta_i\eta_j \middle| \eta_j > 0, \dots, \eta_i > 0\right] \\ &\qquad \times \mathbb{P}\left[\eta_j > 0, \dots, \eta_i > 0\right]. \end{aligned}$$

Due to the independence of the η_i's we have that

1. $$\mathbb{P}\left[\eta_j > 0, \dots, \eta_i > 0\right] = \left(\frac{1}{2}\right)^{i-j+1}$$

   and

2. $$\begin{aligned} \mathbb{E}\left[\eta_i\eta_j \middle| \eta_j > 0, \dots, \eta_i > 0\right] &= \mathbb{E}\left[\eta_i\eta_j \middle| \eta_j > 0, \eta_i > 0\right] \\ &= \mathbb{E}\left[\eta_i \middle| \eta_i > 0\right] \mathbb{E}\left[\eta_j \middle| \eta_j > 0\right]. \end{aligned}$$

Remark 7.2. Let X be a normally distributed random variable and ϕ_X the corresponding probability density function. Then, the density of X conditional on it being positive is

$$\begin{cases} \dfrac{\phi_X(x)}{\mathbb{P}[x > 0]} = 2\phi_X(x), & x > 0 \in X, \\ 0, & \text{otherwise}, \end{cases}$$

since $\mathbb{P}[x > 0] = \dfrac{1}{2}$.

Using Remark 7.2, for a standard normal distributed random variable H we get for every $\eta \in H$ the following equality is true,

$$\mathbb{E}\left[\eta\middle|\eta > 0\right] = 2\int_0^{\infty} \eta\phi_H(\eta)d\eta = \frac{2}{\sqrt{2\pi}}.$$

Thus, $\mathbb{E}\left[\eta_i\middle|\eta_i > 0\right] = \mathbb{E}\left[\eta_j\middle|\eta_j > 0\right] = \frac{2}{2\pi}$.

Putting everything together, one has

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}\left[D_2^{\text{Ext}}\left(W_n\right)_T^2\right] &= T + 4\lim_{n\to\infty}\frac{2T}{n\pi}\sum_{i=2}^{n}\sum_{j=1}^{i-1}\left(\frac{1}{2}\right)^{i-j+1} \\
&= T + 4\lim_{n\to\infty}\frac{T}{n\pi}\sum_{i=2}^{n}\sum_{j=1}^{i-1}\left(\frac{1}{2}\right)^{i-j} \\
&= T\left(1 + \lim_{n\to\infty}\left(\frac{4}{\pi} + \frac{4\cdot 2^{1-n}}{n\pi} - \frac{8}{n\pi}\right)\right) \qquad (7.7)\\
&= T\left(1 + \frac{4}{\pi}\right).
\end{aligned}$$

Similarly, using the expression in Eq.(7.5) one has

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}\left[D_2^{\text{Ext}}\left(W_n\right)_T^2\right] &= T + 4\lim_{n\to\infty}\frac{2T}{n\pi}\sum_{k=2}^{n}(n+1-k)\left(\frac{1}{2}\right)^{i} \\
&= T\left(1 + \frac{4}{\pi}\right).
\end{aligned}$$

□

Table 7.1 presents both the analytical and numerical value of the expectation of the (ExtQV) for the Brownian motion for five values of $n = \left(10^2, 10^3, 10^4, 10^5, 10^6\right)$. The analytical value is given by Eq.(7.7) and the numerical value using the Algorithm 1 for 1000 realizations of Brownian motion paths. As it can be seen from Table 7.1, the absolute difference between the analytical and numerical values is very small and in fact as n increases the absolute difference becomes negligible.

In the following section, we examine the behavior of the (ExtQV) when it is applied to data generated by model 7.1.

N	Analytical	Numerical	Absolute Difference
10^2	2.2478	2.2425	0.0053
10^3	2.2707	2.2751	0.0044
10^4	2.2730	2.2710	0.0020
10^5	2.2732	2.2731	1.38×10^{-4}
10^6	2.2732	2.2732	4.55×10^{-5}

Table 7.1: Theoretical and Numerical expectation of the (ExtQV) for the Brownian Motion for different n's.

7.2 Extrema Quadratic Variation for Multiscale Diffusions with Bounded Quadratic Variation

Consider the following fast/slow system of SDEs

$$dx^\epsilon(t) = \frac{\sigma_1}{\epsilon} y^\epsilon(t)dt + \sigma_2 dU(t), \qquad x^\epsilon(0) = x_0, \tag{7.8a}$$

$$dy^\epsilon(t) = -\frac{1}{\epsilon^2} y^\epsilon(t)dt + \frac{1}{\epsilon} dW(t), \qquad y^\epsilon(0) = y_0, \tag{7.8b}$$

where U and W are independent standard Brownian motions. From Eq.(7.8) it is easy to see that

$$dx^\epsilon(t) = \sigma_1 \left(dW(t) - \epsilon dy^\epsilon(t)\right) + \sigma_2 dU(t), \tag{7.9}$$

and therefore the corresponding homogenized SDE for this model is

$$dX(t) = \sigma_1 dW(t) + \sigma_2 dU(t), \qquad X(0) = x_0. \tag{7.10}$$

In what follows, we examine if the (ExtQV) of the x^ϵ process is asymptotically unbiased to the quadratic variation of the X process which in this case is equal to $(\sigma_1 + \sigma_2)$. In particular, using Algorithm 1, we compute the square of the (ExtQV) for data generated by (7.1) and we examine if by increasing n we can obtain the real value of $(\sigma_1 + \sigma_2)^2$. Table 7.2 illustrates the values of the expectation of the squared (ExtQV) for different values of ϵ and n and suggests that this is not the case. In fact, from the results in Table 7.2 we deduce that the expectation of the (ExtQV) tends to the quantity $\sigma_2^2 T \left(1 + \frac{4}{\pi}\right)$. In other words, the presence of the noise term in Eq.(7.8a) consists the (ExtQV) insufficient in the limit. Probably the reason for this is the fact that the Brownian motion is rougher than the bounded variation model considered it Chapter 4 and its extremals are also the extremals of the x^ϵ process given by Eq.(7.8).

	n			
ϵ	10^4	10^5	10^6	10^7
0.20	2.2793	2.2741	2.3731	2.2732
0.15	2.2922	2.2746	2.2737	2.2733
0.10	2.3084	2.2759	2.2736	2.2733
0.05	2.4266	2.2873	2.2746	2.2734
0.01	4.5419	2.6463	2.3115	2.2771

Table 7.2: Expectation of the (ExtQV) for different ϵ's and n's.

Judging from the results in Table 7.2, we cannot use (ExtQV) to estimate the limiting diffusion coefficient in the limit of $n \to \infty$. However, we can use (ExtQV) to correctly identify the homogenized coefficient for the appropriate combination of ϵ and n. Indeed, Table 7.3 illustrates the value of the (ExtQV) for n and ϵ such that $n\epsilon^2 \approx \sqrt{2}$. In general, through numerical studies we conclude that we should seek for a relation of the form $n\epsilon^\zeta \sim \mathcal{O}(1)$ where ζ in the neighbourhood of 2. Note that this is a different range from the one in Pavliotis and Stuart (2007) where (QV) was applied on subsampled data.

n	$\sigma_1 = 1$,$\sigma_2 = 1$	$\sigma_1 = 1$,$\sigma_2 = 2$	$\sigma_1 = 2$,$\sigma_2 = 2$
10^3	4.2629	9.5815	16.8102
10^4	4.1236	9.2932	16.5981
10^5	4.0384	9.1502	16.2046
10^6	4.0107	9.0402	16.1238
$(\sigma_1 + \sigma_2)^2$	4	9	16

Table 7.3: Expectation of (ExtQV) for different n's, σ_1's, σ_2's and for $\epsilon = \dfrac{2^{1/4}}{\sqrt{n}}$.

7.3 Summary

Initially in this chapter we considered the case of Browinan motion and we proved that the expectation of the (ExtQV) of the standard Brownian motion defined on the interval $[0, T]$ is asymptotically equal to the quantity $T\left(1 + \dfrac{4}{\pi}\right)$. Then, we illustrated by example that this is also true (proportionally) for multiscale models of bounded quadratic variation. Despite the fact that our proposed estimator is not asymptotically unbiased for this class of models, for the appropriate combination of ϵ and n we can identify correctly the homogenized coefficient.

CHAPTER 8

Conclusions and Future Work

This last chapter contains a summary of the thesis, outlines the contributions and gives overall conclusions. Also, we discuss some limitations of our work and open research areas and provide ideas for future work.

8.1 Thesis Summary

In this thesis, we have examined fast/slow systems of SDEs for which a coarse–grained model can be found for their slow dynamics. In reality such models are not explicitly known, and the aim is to fit a SDE to data that assumed to have multi-scale character by estimating the free parameters of the model. We considered data possessing two widely separated time scales and we employed the homogenization case for multiscale diffusions to model such data. Our main objective was to find an efficient estimator for the diffusion coefficient of the homogenized SDE. Throughout this thesis we have mentioned that it is important to construct an estimator that does not depend on the explicit knowledge of the scale separation parameter. Our main contribution has been that our proposed estimator does not require the knowledge of this parameter. The efficiency of our proposed estimator has been demonstrated on a variety of different models exhibiting zero quadratic variation.

In Chapter 2, we briefly reviewed elements from the theory of multiscale diffusions under which a homogenized SDE can be found for such systems. Knowing the form of the homogenized SDE beforehand gave us the benefit of being able to test our estimator numerically.

In Chapter 3, we introduced our proposed estimator, the extrema quadratic variation

(ExtQV). Our proposed estimator suggested in taking the sum of squared returns of the process evaluated at its extremal points. It should be considered as a way of subsampling the data. Subsampling is already the standard approach in literature for such models. However, to find the optimal subsampling rate the knowledge of the scale separation parameter is necessary. The advantage of our estimator is exactly the fact that it does not require the knowledge of it.

In Chapter 4, we applied the estimator to data that exhibit zero quadratic variation. To be more specific, we considered a simple multiscale system were fast dynamics are described by an Ornstein–Uhlenbeck (OU) process and that allowed us to perform explicit calculations for the expectation of the (ExtQV). We proved theoretically that the (ExtQV) is asymptotically unbiased for the diffusion coefficient of the homogenized SDE. Moreover, this was also verified by numerical experiments. Based on numerical results, we have also shown that our proposed estimator performs with an L_2–error of order ϵ^2. In the existing literature, in the context of multiscale diffusions, the L_2–error of the estimators was at best of order ϵ.

In Chapter 5, we extended our work for more general models of zero quadratic variation. Again, under the appropriate technical assumptions, we proved that the (ExtQV) is asymptotically unbiased for the diffusion coefficient of the corresponding homogenized SDE.

In Chapter 6, we demonstrated the performance of our proposed estimator in four particular examples. The first two were an immediate extension of the multiscale Ornstein–Uhlenbeck (OU) considered in Chapter 4 with the difference lying on the function considered in the slow dynamics of the model. The third example was slightly different in the sense that the fast dynamics were not described by an OU process. Finally, in the last example we considered a model whose fast dynamics were not described by an OU process. The main conclusion from these examples is that our proposed estimator is asymptotically unbiased and consistent for a wide range models.

In Chapter 7, multiscale systems of bounded and non–zero quadratic variation has been considered. In particular, we have demonstrated through an example that the expectation of the (ExtQV) fails to go in the limit to the diffusion coefficient. In fact, it was shown, that in the limit of $n \to \infty$, the expectation of the (ExtQV) goes to the (ExtQV) of the noise term that was added to the model. However, we have numerical evidence supporting that given a piecewise linear approximation of the system's slow variables on a partition $\pi_n([0, T])$, then, for the appropriate choice of ϵ such that $n\epsilon^{\zeta}$, ζ in the neighbourhood of 2, one can correctly identify the homogenized diffusion coefficient.

8.2 Future Work

Here possibilities of future work and improvements are highlighted. There is scope to work theoretically on the computation of the L_2–error for the deterministic model considered in Chapter 4. Also, for the bounded quadratic variation models, it is important to justify in theory the combination of n and ϵ for which our estimator works. Furthermore, an extension to more general cases in a methodology based on Chapter 5 would be very interesting and usefull in practical applications.

Bibliography

Abry, P., Baraniuk, R., Flandrin, P., Riedi, R. and Veitch, D. (2002) Multiscale nature of network traffic. *IEEE Signal Processing Magazine*, **19**, 28–46.

Aït-Sahalia, Y. and Jacod, J. (2014) *High-Frequency Financial Econometrics*. Princeton University Press, Princeton., first edn.

Aït-Sahalia, Y., Mykland, P. A. and Zhang, L. (2005) How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial studies*, **18**, 351–416.

— (2011) Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics*, **160**, 160–175.

Ait-Sahalia, Y. and Xiu, D. (2017) A hausman test for the presence of market microstructure noise in high frequency data.

Andersen, T., Bollerslev, T., Diebold, F. and Labys, P. (2000) Great realizations. *Risk*, **13**, 105–108.

Andersen, T. G., Bollerslev, T. and Diebold, F. X. (2002) Parametric and nonparametric volatility measurement.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2003) Modeling and forecasting realized volatility. *Econometrica*, **71**, 579–625.

Awartani, B., Corradi, V. and Distaso, W. (2004) Testing and modelling market microstructure effects with an application to the dow jones industrial average. *WP04-06*.

Azencott, R., Beri, A. and Timofeyev, I. (2010) Adaptive sub-sampling for parametric estimation of gaussian diffusions. *Journal of Statistical Physics*, **139**, 1066–1089.

— (2011) Parametric estimation of stationary stochastic processes under indirect observability. *Journal of Statistical Physics*, **144**, 150–170.

Bandi, F. M. and Russell, J. R. (2008) Microstructure noise, realized variance, and optimal sampling. *The Review of Economic Studies*, **75**, 339–369.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2008) Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, **76**, 1481–1536.

— (2008a) Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, **76**, 1481–1536.

— (2008b) Multivariate realized kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non- synchronous trading. *Technical report*.

— (2009) Realized kernels in practice: trades and quotes. *The Econometrics Journal*, **12**, C1–C32.

— (2011) Subsampling realised kernels. *Journal of Econometrics*, **160**, 204–219.

Barndorff-Nielsen, O. E. and Shephard, N. (2002) Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **64**, 253–280.

Bensoussan, A., Lions, J. and Papanicolaou, G. (2011) *Asymptotic analysis for periodic structures*, vol. 374. American Mathematical Soc.

Billingsley, P. (2008) *Probability and measure*. New York: John Wiley & Sons.

Bishwal, J. P. (2008) *Parameter estimation in stochastic differential equations*. Springer Berlin.

Brown, S. J. (1990) Estimating volatility. *Financial Options: From Theory to Practice*, **516**, 537.

Christensen, K., Oomen, R. C. and Podolskij, M. (2014) Fact or friction: Jumps at ultra high frequency. *Journal of Financial Economics*, **114**, 576–599.

Corsi, F., Zumbach, G., Muller, U. A. and Dacorogna, M. M. (2001) Consistent high-precision volatility from high-frequency data. *Economic Notes*, **30**, 183–204.

Cotter, C., Pavliotis, G. et al. (2009) Estimating eddy diffusivities from noisy lagrangian observations. *Communications in Mathematical Sciences*, **7**, 805–838.

Crommelin, D. (2012) Estimation of space-dependent diffusions and potential landscapes from non-equilibrium data. *Journal of Statistical Physics*, **149**, 220–233.

Friz, P. K. and Victoir, N. B. (2010) *Multidimensional stochastic processes as rough paths: theory and applications*, vol. 120. Cambridge University Press.

Gailus, S. and Spiliopoulos, K. (2017a) Discrete-time statistical inference for multiscale diffusions in the averaging and homogenization regime. *arXiv preprint arXiv:1709.02223*.

— (2017b) Statistical inference for perturbed multiscale dynamical systems. *Stochastic Processes and their Applications*, **127**, 419–448.

Hansen, P. R. and Lunde, A. (2006) Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, **24**, 127–161.

Hautsch, N. and Podolskij, M. (2013) Preaveraging-based estimation of quadratic variation in the presence of noise and jumps: theory, implementation, and empirical evidence. *Journal of Business & Economic Statistics*, **31**, 165–183.

Iacus, S. M. (2009) *Simulation and inference for stochastic differential equations: with R examples.* Springer Science & Business Media.

Imkeller, P., Namachchivaya, N. S., Perkowski, N., Yeong, H. C. et al. (2013) Dimensional reduction in nonlinear filtering: a homogenization approach. *The Annals of Applied Probability*, **23**, 2290–2326.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M. and Vetter, M. (2009) Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic processes and their applications*, **119**, 2249–2276.

Jacod, J. and Protter, P. (2012) *Discretization of Processes.* Springer-Verlag, New York., first edn.

Jacod, J. and Shiryaev, A. (2013) *Limit theorems for stochastic processes*, vol. 288. Springer Science & Business Media.

Kalliadasis, S., Krumscheid, S. and Pavliotis, G. A. (2015) A new framework for extracting coarse-grained models from time series with multiscale structure. *Journal of Computational Physics*, **296**, 314–328.

Karatzas, I. and Shreve, S. (2012) *Brownian motion and stochastic calculus*, vol. 113. Springer Science & Business Media.

Katsoulakis, M., Majda, A. and Sopasakis, A. (2004) Multiscale couplings in prototype hybrid deterministic/stochastic systems: Part i, deterministic closures. *Communications in Mathematical Sciences*, **2**, 255–294.

— (2005) Multiscale couplings in prototype hybrid deterministic/stochastic systems: Part ii, stochastic closures. *Communications in Mathematical Sciences*, **3**, 453–478.

Kloeden, P. E. and Platen, E. (1999) *Numerical solution of stochastic differential equations*. Applications of mathematics. Berlin, New York: Springer.

Krumscheid, S. (2014a) Perturbation-based inference for diffusion processes: Obtaining coarse-grained models from multiscale data. *arXiv preprint arXiv:1409.4685*.

— (2014b) Statistical and numerical methods for diffusion processes with multiple scales.

Krumscheid, S., Pavliotis, G. A. and Kalliadasis, S. (2013) Semiparametric drift and diffusion estimation for multiscale diffusions. *Multiscale Modeling & Simulation*, **11**, 442–473.

Krumscheid, S., Pradas, M., Pavliotis, G. and Kalliadasis, S. (2015) Data-driven coarse graining in action: Modeling and prediction of complex systems. *Physical Review E*, **92**, 042139.

Kutoyants, Y. A. (2013) *Statistical inference for ergodic diffusion processes*. Springer Science & Business Media.

Ladroue, C. and Papavasiliou, A. (2013) A distributed procedure for computing stochastic expansions with mathematica. *Journal of Statistical Software*, **53**.

Liptser, R. S. and Shiryaev, A. N. (2013) *Statistics of random Processes: I. general Theory*, vol. 5. Springer Science & Business Media.

Lyons, T. and Qian, Z. (2002) *System control and rough paths*. Oxford University Press.

Lyons, T. J. (1998) Differential equations driven by rough signals. *Revista Matemática Iberoamericana*, **14**, 215–310.

Majda, A., Timofeyev, I. and Vanden-Eijnden, E. (2006) Stochastic models for selected slow variables in large deterministic systems. *Nonlinearity*, **19**, 769.

Majda, A. J., Timofeyev, I. and Vanden Eijnden, E. (2001) A mathematical framework for stochastic climate models. *Communications on Pure and Applied Mathematics*, **54**, 891–974.

Mykland, P. A. and Zhang, L. (2016) Between data cleaning and inference: Pre-averaging and robust estimators of the efficient price. *Journal of Econometrics*, **194**, 242–262.

Øksendal, B. (2003) Stochastic differential equations. In *Stochastic differential equations*, 65–84. Springer.

Papanicolaou, A. and Spiliopoulos, K. (2014) Filtering the maximum likelihood for multiscale problems. *Multiscale Modeling & Simulation*, **12**, 1193–1229.

— (2017) Dimension reduction in statistical estimation of partially observed multiscale processes. *SIAM/ASA Journal on Uncertainty Quantification*, **5**, 1220–1247.

Papavasiliou, A. (2011) Coarse-grained modeling of multiscale diffusions: the p-variation estimates. In *Stochastic Analysis 2010*, 169–190. Springer.

Papavasiliou, A., Pavliotis, G. and Stuart, A. (2009) Maximum likelihood drift estimation for multiscale diffusions. *Stochastic Processes and their Applications*, **119**, 3173–3210.

Pavliotis, G. A. (2014) *Stochastic processes and applications: Diffusion processes, the Fokker-Planck and Langevin equations*, vol. 60. Springer.

Pavliotis, G. A., Pokern, Y. and Stuart, A. M. (2008) Parameter estimation for multiscale diffusions: an overview. *Statistical methods for stochastic differential equations*, **124**, 429–472.

Pavliotis, G. A. and Stuart, A. (2007) Parameter estimation for multiscale diffusions. *Journal of Statistical Physics*, **127**, 741–781.

— (2008) *Multiscale methods: averaging and homogenization*. New York: Springer Science & Business Media.

Podolskij, M. and Vetter, M. (2009) Bipower-type estimation in a noisy diffusion setting. *Stochastic processes and their applications*, **119**, 2803–2831.

Podolskij, M., Vetter, M. et al. (2009) Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli*, **15**, 634–658.

Schlick, T. (2010) *Molecular modeling and simulation: an interdisciplinary guide: an interdisciplinary guide*, vol. 21. Springer Science & Business Media.

Spiliopoulos, K. and Chronopoulou, A. (2013) Maximum likelihood estimation for small noise multiscale diffusions. *Statistical Inference for Stochastic Processes*, **16**, 237–266.

Sykulski, A., Olhede, S. C. and Pavliotis, G. A. (2008) Multiscale inference for high-frequency data. *arXiv preprint arXiv:0803.0392*.

Tsay, R. S. (2005) *Analysis of financial time series*, vol. 543. John Wiley & Sons.

Williams, D. (1991) *Probability with martingales.* Cambridge university press.

Xiu, D. (2010) Quasi-maximum likelihood estimation of volatility with high frequency data. *Journal of Econometrics*, **159**, 235–250.

Zhang, L., Mykland, P. A. and Aït-Sahalia, Y. (2005) A tale of two time scales: Determining integrated volatility with noisy high–frequency data. *Journal of the American Statistical Association*, **100**, 1344–1411.

Zhang, L. et al. (2006) Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli*, **12**, 1019–1043.

Zhou, B. (1996) High-frequency data and volatility in foreign-exchange rates. *Journal of Business & Economic Statistics*, **14**, 45–52.

APPENDIX A

Useful Tools

Lemma A.1 (Tower Property). *(Williams, 1991, p. 88) If $\mathcal{H}$ is a sub–σ–algebra of $\mathcal{G}$, then*

$$\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{G}\right]|\mathcal{H}\right] = \mathbb{E}\left[X|\mathcal{H}\right]. \tag{A.1}$$

Lemma A.2 ('Taking out what is known (TOWIK)'). *(Williams, 1991, p. 88) If Z is $\mathcal{G}$–measurable and bounded, then*

$$\mathbb{E}\left[ZX|\mathcal{G}\right] = Z\mathbb{E}\left[X|\mathcal{G}\right], \qquad a.s. \tag{A.2}$$

Lemma A.3 (Jensen's Inequality). *(Billingsley (2008, p. 283)) If f is convex then*

$$\mathbb{E}\left[f(X)\right] \geq f\left(\mathbb{E}\left[X\right]\right). \tag{A.3}$$

If f is concave then

$$\mathbb{E}\left[f(X)\right] \leq f\left(\mathbb{E}\left[X\right]\right). \tag{A.4}$$

Lemma A.4 (Cauchy–Schwarz Inequality). *(Billingsley (2008, p. 283)) Suppose $X, Y \in L^2(\omega, \mathcal{G}, \mathbb{P})$, and suppose $\mathcal{G} \subset \mathcal{F}$ is a σ–algebra. Then*

$$\mathbb{E}\left[XY|\mathcal{G}\right] \leq \left(\mathbb{E}\left[X^2|\mathcal{G}\right]\right)^{1/2}\left(\mathbb{E}\left[Y^2|\mathcal{G}\right]\right)^{1/2}, \ a.s. \tag{A.5}$$

Theorem A.5 (Bounded Convergence Theorem (BCT)). *(Williams (1991)) Suppose that $X_n \to X$ a.s. as $n \to \infty$ and there exists $c \in \mathbb{R}$ such that*

$$|X_n(\omega)| \leq c \leq \infty \tag{A.6}$$

for $n \geq 0$ and $\omega \in \Omega$. Then $\mathbb{E}[x_n] \to \mathbb{E}[X]$ as $n \to \infty$.

Theorem A.6 (Itô formula). *(Pavliotis and Stuart (2008, p. 89)) Let $z(t)$ solve*

$$dz(t) = h(z)dt + \gamma(z)dW(t), \qquad z(0) = z_0.$$

where both $h(\cdot)$ and $\gamma(\cdot)$ are globally Lipschitz on $\mathcal{Z}$ and z_0 is a random variable independent of the Brownian motion $W(t)$ with $\mathbb{E}|z_0|^2 < \infty$. Let $f \in C^2(\mathcal{Z}, \mathbb{R})$, then the process $f(z(t))$ satisfies

$$f(z(t)) = f(z_0) + \int_0^t \mathcal{L}_0 f(z(s))ds + \int_0^t \nabla f(z(s))dW(s),$$

where $\mathcal{L}_0 = h(z)\nabla + \frac{1}{2}\gamma(z)\gamma(z)^{\mathrm{T}}\nabla\nabla$.

Lemma A.7 (Itô isometry). *(Williams (1991)) Let W the canonical real–valued Brownian motion and $X : [0, T] \times \Omega \to \mathbb{R}$ be a stochastic process that is adapted to the natural filtration of the Brownian motion. Then,*

$$\mathbb{E}\left[\left(\int_0^T X(t)dW(t)\right)^2\right] = \mathbb{E}\left[\int_S^T X(t)^2 dt\right].$$

APPENDIX B

Appendix for Chapter 4

Lemma B.1. *Let* $a := e^{-\frac{1}{n\epsilon^2}}$. *Then,*

$$\epsilon^2(1-a)^2 \sum_{k=2}^{n}(n+1-k)a^{k-1} \to 1 + \epsilon^2(e^{-1/\epsilon^2} - 1), \tag{B.1}$$

as $n \to \infty$.

Proof.

$$\begin{aligned} \sum_{k=2}^{n}(n+1-k)a^{k-1} &= \sum_{k=1}^{n-1}(n-k)a^k \\ &= n\sum_{k=1}^{n-1}a^k - \sum_{k=1}^{n-1}ka^k \end{aligned} \tag{B.2}$$

Making use of the the geometric series, i.e.,

$$\sum_{k=0}^{n}a^k = \frac{1-a^{n+1}}{1-a}, \text{ and } \sum_{k=1}^{n}ka^k = \frac{a(1-(n+1)a^n + na^{n+1})}{(1-a)^2}, \tag{B.3}$$

we get that

$$\begin{aligned} \text{Eq.(B.2)} &= n\left(\frac{1-a^n}{1-a} - 1\right) - \frac{a(1-na^{n-1}+(n-1)a^n)}{(1-a)^2} \\ &= n\left(\frac{1-a^n-1+a}{1-a} - 1\right) - \frac{(a-na^n+naa^n-aa^n)}{(1-a)^2} \\ &= \frac{n(a-a^n)(1-a) - na^n(a-1) - a(1-a^2)}{(1-a)^2} \end{aligned}$$

$$= \frac{-a(1-a^n-n+na)}{(1-a)^2}. \tag{B.4}$$

Thus,

$$(1-a)^2 \sum_{k=2}^{n} (n+1-k) a^{k-1} = -a(1-a^n-n+na), \tag{B.5}$$

and since $a^n = e^{-1/\epsilon^2}$ we get

$$\epsilon^2 \lim_{n\to\infty} (1-a)^2 \sum_{k=2}^{n} (n+1-k) a^{k-1} = 1 + \epsilon^2 \left(e^{-1/\epsilon^2} - 1 \right), \tag{B.6}$$

as required. □

Lemma B.2. *Let*

$$M_i := \frac{\int_{t_i}^{t_{i+1}} \left(e^{-\frac{(t_{i+1}-u)}{\epsilon^2}} - 1 \right) dW(u)}{\epsilon (1-a)}, \quad t_i = i\delta = i\frac{T}{n}, a = e^{-\frac{\delta}{\epsilon^2}}.$$

Then,

$$Var[M_i] = \mathbb{E}\left[M_i^2\right] = \frac{\delta + \frac{\epsilon^2}{2}\left(4e^{-\frac{\delta}{\epsilon^2}} - e^{-\frac{2\delta}{\epsilon^2}} - 3\right)}{\epsilon^2 \left(1 - e^{-\frac{\delta}{\epsilon^2}}\right)^2} \approx \frac{\frac{\delta^3}{3\epsilon^4} + \mathcal{O}\left(\frac{\delta^4}{\epsilon^6}\right)}{\epsilon^2 (1-a)^2}. \tag{B.7}$$

Proof.

$$\mathbb{E}\left[M_i^2\right] = \frac{1}{\epsilon^2 (1-a)^2} \mathbb{E}\left[\left(\int_{t_i}^{t_{i+1}} \left(e^{-\frac{(t_{i+1}-u)}{\epsilon^2}} - 1 \right) dW(u) \right)^2 \right], \tag{B.8}$$

which by Itô isometry (see Lemma A.7) is equal to

$$\begin{aligned} \mathbb{E}\left[M_i^2\right] &= \frac{1}{\epsilon^2 (1-a)^2} \mathbb{E}\left[\int_{t_i}^{t_{i+1}} \left(e^{-\frac{(t_{i+1}-u)}{\epsilon^2}} - 1 \right)^2 du \right] \\ &= \frac{\delta + \frac{\epsilon^2}{2}\left(4e^{-\frac{\delta}{\epsilon^2}} - e^{-\frac{2\delta}{\epsilon^2}} - 3\right)}{\epsilon^2 (1-a)^2} \approx \frac{\frac{\delta^3}{3\epsilon^4} + \mathcal{O}\left(\frac{\delta^4}{\epsilon^6}\right)}{\epsilon^2 (1-a)^2}. \end{aligned}$$

as required. □

Lemma B.3. *Let*

$$K_i := \frac{1}{\epsilon} \int_{t_i}^{t_{i+1}} e^{-\frac{(t_{i+1}-u)}{\epsilon^2}} dW(u).$$

Then,

$$Var\left[K_i\right] = \mathbb{E}\left[K_i^2\right] = \frac{1-a^2}{2}. \tag{B.9}$$

Proof. Similarly to Lemma B.2

$$\begin{aligned}
\mathbb{E}\left[K_i^2\right] &= \frac{1}{\epsilon}\mathbb{E}\left[\left(\int_{t_i}^{t_{i+1}} e^{-\frac{(t_{i+1}-u)}{\epsilon^2}} dW(u)\right)^2\right] \\
&= \frac{1}{\epsilon^2}\mathbb{E}\left[\int_{t_i}^{t_{i+1}} e^{-\frac{2(t_{i+1}-u)}{\epsilon^2}} du\right] \\
&= \frac{1-a^2}{2}
\end{aligned}$$

as required. □

Lemma B.4. *Let Φ_X and ϕ_X be the cdf and the pdf of a symmetric random variable X with mean zero and variance σ_X^2. Then, the following inequality holds*

$$\sigma_X^2\phi_X(x) - x\Phi_X(x) \leq \begin{cases} \sigma_X^2\phi_X(x), & x<0 \\ \sigma_X^2\phi_X(x) - x, & x>0 \end{cases}. \tag{B.10}$$

Proof. Let $x>0$, then since $0 \leq \Phi_X(x) \leq 1$ one has $x\Phi_X(x) \geq 0$ and therefore

$$\sigma_X^2\phi_X(x) - x\Phi_X(x) \leq \sigma_X^2\phi_X(x).$$

For $x<0$, the fact that $0 \leq \Phi_X(x) \leq 1$ implies that $-x\Phi_X(x) \leq -x$ and as before

$$\sigma_X^2\phi_X(x) - x\Phi_X(x) \leq \sigma_X^2\phi_X(x) - x$$

as required. □

Lemma B.5. *Show that* $\mathbb{P}\left[y_n^\epsilon(t_1) > 0 \middle| y_n^\epsilon(t_0) > 0\right] = 1 - \dfrac{\arctan\left(\frac{\sqrt{1-a^2}}{a}\right)}{\pi}$.

Proof.

$$\begin{aligned}
\mathbb{P}\left[y_n^\epsilon(t_1) > 0 \middle| y_n^\epsilon(t_0) > 0\right] &= \frac{\mathbb{P}\left[y_n^\epsilon(t_1) > 0, y_n^\epsilon(t_0) > 0\right]}{\mathbb{P}\left[y_n^\epsilon(t_0) > 0\right]} \\
&= 2\mathbb{P}\left[y_n^\epsilon(t_1) > 0, y_n^\epsilon(t_0) > 0\right] \\
&\overset{\text{Eq.(4.24)}}{=} 2\mathbb{P}\left(K_{t_0} > -a y_n^\epsilon(t_0), y_n^\epsilon(t_0) > 0\right).
\end{aligned} \tag{B.11}$$

Now, let $\Phi_K, \phi_{y_n^\epsilon}$ be the cdf of K and the pdf of $y_n^\epsilon(t_0)$ respectively, namely the normal cdf with mean zero and variance σ_K^2 and the normal with mean zero and variance $\frac{1}{2}$ respectively. Then, the probability in Eq.(B.11) is equal to

$$\begin{aligned}
\mathbb{P}\left[y_n^\epsilon(t_1) > 0 \middle| y_n^\epsilon(t_0) > 0\right] &= 2\int_0^\infty \left(1 - \Phi_K(-a y_n^\epsilon(t_0))\right) \phi_{y_n^\epsilon}(y_n^\epsilon(t_0)) dy_n^\epsilon(t_0) \\
&= 1 - \frac{\arctan\left(\frac{\sqrt{1-a^2}}{a}\right)}{\pi},
\end{aligned}$$

as required. □

Lemma B.6. *Let* $a = e^{-\frac{\delta}{\epsilon^2}}$ *and* $\rho'_{n,\epsilon} = \left(1 - \frac{\arctan \frac{\sqrt{1-a^2}}{a}}{\pi}\right)^{1/2}$. *Then,*

$$\ell = \frac{2\sigma^2\epsilon^2}{\sqrt{2\pi}} \lim_{n\to\infty} (1-a)^2 \frac{(1-a^2)^{3/4}}{(1+a^2)^{1/4}} \sum_{k=2}^{n} (n+1-k) \sum_{j=0}^{k-2} a^j \left(\rho'_{n,\epsilon}\right)^{k-2-j} = 0.$$

Proof.

$$\begin{aligned}
\ell &= \frac{2\sigma^2\epsilon^2}{\sqrt{2\pi}} \lim_{n\to\infty} (1-a)^2 \frac{(1-a^2)^{3/4}}{(1+a^2)^{1/4}} \sum_{k=2}^{n} (n+1-k) \left(\rho'_{n,\epsilon}\right)^{k-2} \sum_{j=0}^{k-2} \left(\frac{a}{\rho'_{n,\epsilon}}\right)^j \\
&= \frac{2\sigma^2\epsilon^2}{\sqrt{2\pi}} \lim_{n\to\infty} \frac{(1-a)^2}{\left(\rho'_{n,\epsilon} - a\right)} \frac{(1-a^2)^{3/4}}{(1+a^2)^{1/4}} \sum_{k=2}^{n} (n+1-k) \left(\left(\rho'_{n,\epsilon}\right)^k - a^{k-1}\right) \\
&= \frac{2\sigma^2\epsilon^2}{\sqrt{2\pi}} \lim_{n\to\infty} \frac{(1-a)^2}{\left(\rho'_{n,\epsilon} - a\right)} \frac{(1-a^2)^{3/4}}{(1+a^2)^{1/4}} \sum_{k=2}^{n} (n+1-k) \left(\rho'_{n,\epsilon}\right)^k \qquad \text{(B.12a)} \\
&\quad - \frac{2\sigma^2\epsilon^2}{\sqrt{2\pi}} \lim_{n\to\infty} \frac{(1-a)^2}{\left(\rho'_{n,\epsilon} - a\right)} \frac{(1-a^2)^{3/4}}{(1+a^2)^{1/4}} \sum_{k=2}^{n} (n+1-k) a^{k-1}. \qquad \text{(B.12b)}
\end{aligned}$$

For the second part (Eq.(B.12b)), first notice that

$$\sum_{k=2}^{n} (n+1-k) a^{k-1} = \frac{a\left[(1-a^n) - n(1-a)\right]}{(1-a)^2}. \qquad \text{(B.13)}$$

Thus, Eq.(B.12b) becomes

$$\begin{aligned}
\text{Eq.(B.12b)} &= \frac{2\sigma^2\epsilon^2}{\sqrt{2\pi}} \lim_{n\to\infty} \frac{(1-a^2)^{3/4}}{(1+a^2)^{1/4}} \frac{a\left[(1-a^n) - n(1-a)\right]}{\left(\rho'_{n,\epsilon} - a\right)} \\
&= \frac{2\sigma^2\epsilon^2}{\sqrt{2\pi}} \lim_{n\to\infty} \frac{(1-a^2)^{3/4}}{(1+a^2)^{1/4}} \frac{a\left(1-a^n\right)}{\left(\rho'_{n,\epsilon} - a\right)} \qquad \text{(B.14a)}
\end{aligned}$$

$$-\frac{2\sigma^2\epsilon^2}{\sqrt{2\pi}}\lim_{n\to\infty}\frac{(1-a^2)^{3/4}}{(1+a^2)^{1/4}}\frac{an(1-a)}{\left(\rho'_{n,\epsilon}-a\right)} \tag{B.14b}$$

Next, notice that $1-a^n = 1-e^{-\frac{1}{\epsilon^2}}$ and that a Taylor series expansion around $\delta = 0$ suggest that

- $\rho'_{n,\epsilon} - a \approx -\frac{1}{\sqrt{2\pi}}\sqrt{\frac{\delta}{\epsilon^2}}$, $1-a \approx \frac{\delta}{\epsilon^2}$, $1-a^2 \approx \frac{2\delta}{\epsilon^2}$, $1+a^2 \approx 2$.
- $1-\rho'_{n,\epsilon} \approx \frac{1}{\sqrt{2\pi}}\sqrt{\frac{\delta}{\epsilon^2}}$, $1-\left(\rho'_{n,\epsilon}\right)^n \approx 1$.

Given the above and by ignoring the constants and the higher order terms of $\frac{\delta}{\epsilon^2}$ we get for Eq.(B.14a)

$$\epsilon^2\lim_{n\to\infty}\frac{(1-a^2)^{3/4}}{(1+a^2)^{1/4}}\frac{a\,(1-a^n)}{\left(\rho'_{n,\epsilon}-a\right)} = \left(\epsilon^2\right)^{3/4}\left(1-e^{-\frac{1}{\epsilon^2}}\right)\lim_{n\to\infty}n^{-1/4} = 0. \tag{B.15}$$

and for Eq.(B.14b)

$$\epsilon^2\lim_{n\to\infty}\frac{(1-a^2)^{3/4}}{(1+a^2)^{1/4}}\frac{an\,(1-a)}{\left(\rho'_{n,\epsilon}-a\right)} = \left(\epsilon^2\right)^{1/4}\lim_{n\to\infty}n^{-1/4} = 0. \tag{B.16}$$

Given Eq.(B.15) and Eq.(B.16) we deduce that Eq.(B.12b) is equal to zero. Finally, for Eq.(B.12a)

$$\begin{aligned}
\text{Eq.(B.12a)} &= \frac{2\sigma^2\epsilon^2}{\sqrt{2\pi}}\lim_{n\to\infty}\frac{(1-a)^2(1-a^2)^{3/4}}{(1+a^2)^{1/4}}\frac{\left(\rho'_{n,\epsilon}\right)^2\left[\left(1-\left(\rho'_{n,\epsilon}\right)^n\right)-n(1-\rho'_{n,\epsilon})\right]}{(1-\rho'_{n,\epsilon})^2\left(\rho'_{n,\epsilon}-a\right)} \\
&= \frac{2\sigma^2\epsilon^2}{\sqrt{2\pi}}\lim_{n\to\infty}\frac{(1-a)^2(1-a^2)^{3/4}}{(1+a^2)^{1/4}}\frac{\left(\rho'_{n,\epsilon}\right)^2\left(1-\left(\rho'_{n,\epsilon}\right)^n\right)}{(1-\rho'_{n,\epsilon})^2\left(\rho'_{n,\epsilon}-a\right)} \qquad \text{(B.17a)} \\
&\quad -\frac{2\sigma^2\epsilon^2}{\sqrt{2\pi}}\lim_{n\to\infty}\frac{(1-a)^2(1-a^2)^{3/4}}{(1+a^2)^{1/4}}\frac{\left(\rho'_{n,\epsilon}\right)^2 n(1-\rho'_{n,\epsilon})}{(1-\rho'_{n,\epsilon})^2\left(\rho'_{n,\epsilon}-a\right)} \qquad \text{(B.17b)}
\end{aligned}$$

$$\text{Eq.(B.17a)} = \left(\epsilon^2\right)^{-1/4}\lim_{n\to\infty}n^{-5/4} = 0 \tag{B.18}$$

and

$$\text{Eq.(B.17b)} = \left(\epsilon^2\right)^{5/4}\lim_{n\to\infty}n^{-3/4} = 0. \tag{B.19}$$

Given Eq.(B.18) and Eq.(B.19) we deduce that Eq.(B.12a) is also equal to zero which concludes the proof. □

B.1 More Numerical Results for Chapter 4

Tables B.1 and B.2 show the expectation of the extrema quadratic vatiation and its corresponding L_2–error for data generated by model (4.1) for different values of the scale separation parameter ϵ and for different time steps $\delta = T/n$. The results in Table B.1 correspond to $\sigma = 1$ and $\sigma = 2$ and the results in Table B.2 correspond to $\sigma = 3$ and $\sigma = 4$.

$\epsilon = 0.01$	$\hat{\sigma}^2$ $(\sigma = 1)$		$\hat{\sigma}^2$ $(\sigma = 2)$	
Sample Size	$\mathbb{E}[\hat{\sigma}^2]$	$L_2[\hat{\sigma}^2]$	$\mathbb{E}[\hat{\sigma}^2]$	$L_2[\hat{\sigma}^2]$
10^3	NaN	NaN	NaN	NaN
10^4	2.2722	1.6209	9.0822	25.8668
10^5	1.2189	0.0494	4.8691	0.7783
10^6	1.0646	0.0053	4.2662	0.0904
10^7	1.0206	0.0016	4.0859	0.0240
$\epsilon = 0.05$			$L_2(\epsilon)/\sigma^4 = 0.0015$	
10^3	1.4962	0.2877	6.0186	4.8063
10^4	1.1278	0.0478	4.5303	0.8421
10^5	1.0419	0.0312	4.1671	0.4706
10^6	1.0077	0.0274	4.0597	0.4166
10^7	1.0062	0.0271	4.0235	0.4515
$\epsilon = 0.10$			$L_2(\epsilon)/\sigma^4 = 0.0282$	
10^3	1.1843	0.1747	4.8497	2.9082
10^4	1.0778	0.1386	4.2536	1.9852
10^5	1.0058	0.1049	4.0037	1.6694
10^6	0.9829	0.1096	3.9876	1.6087
10^7	0.9997	0.1030	3.9098	1.5779
$\epsilon = 0.15$			$L_2(\epsilon)/\sigma^4 = 0.099$	
10^3	1.1153	0.3100	4.4207	4.8360
10^4	1.0145	0.2403	4.0542	3.7374
10^5	0.9837	0.2324	3.8971	3.5146
10^6	0.9582	0.2021	4.0345	3.8575
10^7	0.9523	0.2028	3.8371	3.3327
$\epsilon = 0.20$			$L_2(\epsilon)/\sigma^4 = 0.2083$	
10^3	1.0303	0.4486	4.3331	8.3510
10^4	0.9776	0.3521	3.8314	6.6024
10^5	0.9622	0.4208	3.8308	5.5690
10^6	0.9542	0.3820	3.6215	5.2784
10^7	0.9573	0.3820	3.7935	6.9338
			$L_2(\epsilon)/\sigma^4 = 0.4333$	

Table B.1: Expectation and L_2–error of the (ExtQV) for the model (4.1) for different ϵ's, n's and for $\sigma = 1$ and $\sigma = 2$.

$\epsilon = 0.01$	$\hat{\sigma}^2$ $(\sigma = 3)$		$\hat{\sigma}^2$ $(\sigma = 4)$	
Sample Size	$\mathbb{E}[\hat{\sigma}^2]$	$L_2[\hat{\sigma}^2]$	$\mathbb{E}[\hat{\sigma}^2]$	$L_2[\hat{\sigma}^2]$
10^3	NaN	NaN	NaN	NaN
10^4	20.177	0.1907	36.371	415.63
10^5	10.967	0.1129	19.475	12.4530
10^6	9.5809	0.0990	17.034	1.3624
10^7	9.1933	0.0892	16.338	0.4139
$\epsilon = 0.05$	$L_2(\epsilon)/\sigma^4 = 0.0016$		$L_2(\epsilon)/\sigma^4 = 0.0016$	
10^3	13.597	4.0199	24.064	75.506
10^4	10.172	2.6639	18.089	12.536
10^5	9.3323	2.2839	16.52	7.6939
10^6	9.0713	2.2955	16.13	6.8585
10^7	8.9201	2.0926	15.917	6.2955
$\epsilon = 0.10$	$L_2(\epsilon)/\sigma^4 = 0.0258$		$L_2(\epsilon)/\sigma^4 = 0.0245$	
10^3	10.576	13.101	19.006	35.922
10^4	9.3811	9.2351	16.776	28.472
10^5	9.10001	8.2229	15.904	25.541
10^6	8.8607	7.9155	15.699	24.274
10^7	8.9042	8.9979	15.858	25.589
$\epsilon = 0.15$	$L_2(\epsilon)/\sigma^4 = 0.1110$		$L_2(\epsilon)/\sigma^4 = 0.099$	
10^3	10.026	21.626	17.336	64.95
10^4	9.3327	20.374	15.864	58.024
10^5	8.7393	18.907	15.851	60.544
10^6	8.8361	15.873	15.517	57.739
10^7	8.3540	15.283	15.473	51.121
$\epsilon = 0.20$	$L_2(\epsilon)/\sigma^4 = 0.1936$		$L_2(\epsilon)/\sigma^4 = 0.1997$	
10^3	9.4737	34.398	16.693	110.59
10^4	8.6042	30.976	15.414	96.321
10^5	8.5914	29.89	15.164	94.386
10^6	8.5448	33.412	14.774	85.213
10^7	8.5939	32.862	14.791	87.613
$L_2(\epsilon)/\sigma^4 = 0.4057$			$L_2(\epsilon)/\sigma^4 = 0.3422$	

Table B.2: Expectation and L_2–error of the (ExtQV) for the model (4.1) for different ϵ's, n's and for $\sigma = 3$ and $\sigma = 4$.

APPENDIX C

Appendix for Chapter 5

At this part of the thesis we prove the necessary results that were used in Chapter 5. In what follows, let the functions f, g, β, Φ, the process $y_t := y^{\epsilon}(t)$ and the operator $\mathcal{L}_0$ as defined in Chapter 5.

Definition C.1. *Let $\lambda \in \mathbb{R}$, then*

$$\begin{aligned} e^{\lambda \mathcal{L}_0} f(y) : &= \lim_{N \to \infty} \sum_{m=0}^{N} \frac{\lambda^m}{m!} \mathcal{L}_0^{(m)} f(y), \\ &= \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} \mathcal{L}_0^{(m)} f(y), \end{aligned}$$

which is well defined in the L^{∞} topology by our assumption for the model in Eq.(5.1).

Lemma C.2. *By Itô Lemma (see Theorem A.6) we get*

$$f(y_t) = f(y_0) + \frac{1}{\epsilon^2} \int_0^t \mathcal{L}_0 f(y_{\tau_1}) d\tau_1 + \frac{1}{\epsilon} \int_0^t N^{(f)}(y_{\tau_1}) dV_{\tau_1}. \tag{C.1}$$

Apply Itô–Taylor stochastic expansion (see Kloeden and Platen (1999, Chapter 5)) to obtain that an approximation to the solution of Eq.(C.1) *is given by*

$$f(y_t) = e^{\frac{t}{\epsilon^2} \mathcal{L}_0} f(y_0) + \frac{1}{\epsilon} \int_0^t e^{\frac{(t-s)}{\epsilon^2} \mathcal{L}_0} \left(\nabla_y f \beta \right)(y_s) dV_s. \tag{C.2}$$

Proof. Let $N^{(f)}(y) := \left(\nabla_y f \beta \right)(y)$. Applying Itô's formula to the function $\mathcal{L}_0 f$ this time, we get

$$\mathcal{L}_0 f(y_{\tau_1}) = \mathcal{L}_0 f(y_0) + \frac{1}{\epsilon^2} \int_0^{\tau_1} \mathcal{L}_0^{(2)} f(y_{\tau_2}) d\tau_2 + \frac{1}{\epsilon} \int_0^{\tau_1} N^{(\mathcal{L}_0 f)}(y_{\tau_2}) dV_{\tau_2}, \tag{C.3}$$

where $N^{(\mathcal{L}_0 f)}(y) = \left(\nabla_y \left(\mathcal{L}_0 f \beta\right)\right)(y)$. Substituting Eq.(C.3) into Eq.(C.1) we obtain

$$\begin{aligned} f(y_t) &= f(y_0) + \frac{t}{\epsilon^2}\mathcal{L}_0 f(y_0) + \left(\frac{1}{\epsilon^2}\right)^2 \int_0^t \int_0^{\tau_1} \mathcal{L}_0^{(2)} f(y_{\tau_2}) d\tau_2 d\tau_1 \\ &+ \frac{1}{\epsilon}\left(\frac{1}{\epsilon^2}\right) \int_0^t \int_0^{\tau_1} N^{(\mathcal{L}_0 f)}(y_{\tau_2}) dV_{\tau_2} d\tau_1 + \frac{1}{\epsilon}\int_0^t N^{(f)}(y_{\tau_1}) dV_{\tau_1}. \end{aligned}$$

Similarly to Eq.(C.3),

$$\begin{aligned} \mathcal{L}_0^{(2)} f(y_{\tau_2}) &= \mathcal{L}_0^{(2)} f(y_0) + \frac{1}{\epsilon^2}\int_0^{\tau_2} \mathcal{L}_0^{(3)} f(y_{\tau_3}) d\tau_3 \\ &\qquad + \frac{1}{\epsilon}\int_0^{\tau_2} N^{\left(\mathcal{L}_0^{(2)} f\right)}(y_{\tau_3}) dV_{\tau_3}, \end{aligned}$$

and consequently,

$$\begin{aligned} f(y_t) &= f(y_0) + \frac{t}{\epsilon^2}\mathcal{L}_0 f(y_0) + \frac{1}{2}\left(\frac{t}{\epsilon^2}\right)^2 \mathcal{L}_0^{(2)} f(y_0) \\ &+ \int_0^t \int_0^{\tau_1} \int_0^{\tau_2} \mathcal{L}_0^{(3)} f(y_{\tau_3}) d\tau_3 d\tau_2 d\tau_1 \\ &\quad + \frac{1}{\epsilon}\left(\frac{1}{\epsilon^2}\right)^2 \int_0^t \int_0^{\tau_1} \int_0^{\tau_2} N^{(\mathcal{L}_0^{(2)} f)}(y_{\tau_3}) dV_{\tau_3} d\tau_2 d\tau_1 \\ &\quad\quad + \frac{1}{\epsilon}\left(\frac{1}{\epsilon^2}\right) \int_0^t \int_0^{\tau_1} N^{(\mathcal{L}_0 f)}(y_{\tau_2}) dV_{\tau_2} d\tau_1 \\ &\quad\quad\quad + \frac{1}{\epsilon}\int_0^t N^{(f)}(y_{\tau_1}) dV\tau_1. \end{aligned}$$

Inductively, after r iterations one has

$$\begin{aligned} f(y_t) &= \sum_{m=0}^{r} \left(\frac{t}{\epsilon^2}\right)^m \frac{\mathcal{L}_0^{(m)} f(y_0)}{m!} \\ &+ \int_0^t \left(\frac{1}{\epsilon^2}\right)^r \int_0^{\tau_1} \cdots \int_0^{\tau_r} \mathcal{L}_0^{(r+1)} f(y_{\tau_{r+1}}) d\tau_{r+1} d\tau_r \cdots d\tau_1 \\ &+ \frac{1}{\epsilon}\int_0^t \sum_{m=0}^{r} \left(\frac{1}{\epsilon^2}\right)^m \int_0^{\tau_1} \cdots \int_0^{\tau_m} N^{\left(\mathcal{L}_0^{(m)} f\right)}(y_{\tau_{m+1}}) dV_{\tau_{m+1}} d\tau_m \cdots d\tau_1. \end{aligned} \tag{C.4}$$

Letting $r \to \infty$, then by Definition C.1 the first term of Eq.(C.4) tends to

$$\sum_{m=0}^{r} \left(\frac{t}{\epsilon^2}\right)^m \frac{\mathcal{L}_0^{(m)} f(y_0)}{m!} \to e^{\frac{t}{\epsilon^2}\mathcal{L}_0} f(y_0). \tag{C.5}$$

Also, since the functions $\mathcal{L}^{(r)}f$ are continuous for every r, Fubini's Theorem is applied and by changing the order of integration the second term can be written as

$$\int_0^t \left(\frac{1}{\epsilon^2}\right)^r \frac{(t-\tau_{r+1})^r}{r!} \mathcal{L}_0^{(r+1)} f(y_{\tau_{r+1}}) d\tau_{r+1}.$$

Similarly, for the third term,

$$\begin{aligned} T_3 &= \frac{1}{\epsilon}\int_0^t \sum_{m=0}^{\infty} \left(\frac{1}{\epsilon^2}\right)^m \frac{(t-\tau_{m+1})^m}{m!} N^{\left(\mathcal{L}_0^{(m)}f\right)}(y(\tau_{m+1}))dV(\tau_{m+1}) \\ &= \frac{1}{\epsilon}\int_0^t \sum_{m=0}^{\infty} \left(\frac{1}{\epsilon^2}\right)^m \frac{(t-\tau_{m+1})^m}{m!} \left(\nabla_y \mathcal{L}_0^{(m)} f\beta\right)(y(\tau_{m+1}))dV(\tau_{m+1}) \\ &= \frac{1}{\epsilon}\int_0^t e^{\frac{(t-\tau_{m+1})\mathcal{L}_0}{\epsilon^2}} \left(\nabla_y f\beta\right)(y(\tau_{m+1}))dV(\tau_{m+1}). \end{aligned}$$

However,

$$\lim_{r\to\infty} \left(\frac{1}{\epsilon^2}\right)^r \frac{(t-\tau_{r+1})^r}{r!} = 0,$$

and since the function f and all its derivatives are bounded we get that

$$\lim_{r\to\infty} \int_0^t \left(\frac{1}{\epsilon^2}\right)^r \int_0^{\tau_1} \cdots \int_0^{\tau_r} \mathcal{L}_0^{(r+1)} f(y_{\tau_{r+1}}) d\tau_{r+1} d\tau_r \cdots d\tau_1 = 0.$$

Putting everything together, an approximation of the solution f is given by

$$f(y_t) = e^{\frac{t}{\epsilon^2}\mathcal{L}_0} f(y_0) + \frac{1}{\epsilon}\int_0^t e^{\frac{(t-s)}{\epsilon^2}\mathcal{L}_0} \left(\nabla_y f\beta\right)(y_s)dV_s, \tag{C.6}$$

as required in Eq,(C.2). □

Lemma C.3. *Show that*

$$\frac{1}{\epsilon}\int_{t_{i-1}}^{t_i} e^{\frac{(t-t_{i-1})}{\epsilon^2}\mathcal{L}_0} f(y_{t_{i-1}})dt = \epsilon\left(e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} - 1\right)\mathcal{L}_0^{-1} f(y_{t_{i-1}}),$$

where $\mathcal{L}_0^{-1}$ is the inverse of the operator $\mathcal{L}_0$.

Proof. From Definition C.1 and the fact that $t_i := i\delta$ (recall that for simplicity we consider the case of $T=1$) we get

$$\frac{1}{\epsilon}\int_{t_{i-1}}^{t_i} e^{\frac{(t-(i-1)\delta)}{\epsilon^2}\mathcal{L}_0} f(y(t_{i-1}))dt = \frac{1}{\epsilon}\int_{(i-1)\delta}^{i\delta} \sum_{m=0}^{\infty} \left(\frac{(t-(i-1)\delta)}{\epsilon^2}\right)^m \frac{\mathcal{L}_0^{(m)} f(y_{t_{i-1}})}{m!}$$

$$
\begin{aligned}
&= \frac{1}{\epsilon}\sum_{m=0}^{\infty}\left(\frac{1}{\epsilon^2}\right)^m \frac{\mathcal{L}_0^{(m)} f(y_{t_{i-1}})}{m!}\int_{(i-1)\delta}^{i\delta}(t-(i-1)\delta)^m \\
&= \frac{1}{\epsilon}\sum_{m=0}^{\infty}\left(\frac{1}{\epsilon^2}\right)^m \frac{\delta^{m+1}}{m+1}\frac{\mathcal{L}_0^{(m)} f(y_{t_{i-1}})}{m!} \\
&= \epsilon\sum_{m=0}^{\infty}\left(\frac{\delta}{\epsilon^2}\right)^{m+1} \frac{\mathcal{L}_0^{(m)} f(y_{t_{i-1}})}{(m+1)!} \\
&= \epsilon\left(e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0}-1\right)\mathcal{L}_0^{-1} f(y_{t_{i-1}})
\end{aligned}
$$

as required. □

Lemma C.4. *Let*

$$
M_{t_i} = \frac{1}{\epsilon^2}\int_{t_{i-1}}^{t_i}\int_u^{t_i} e^{\frac{(t-u)}{\epsilon^2}\mathcal{L}_0}(\nabla_y f\beta)(y_u)dtdV_u.
$$

Show that

$$
\left(\mathbb{E}\left[(M_{t_i})^2\right]\right)^{1/2} \leq \frac{C}{\sqrt{3}n^{3/2}\epsilon^{7/2}},
$$

where $C \in \mathbb{R}$.

Proof. From Definition C.1 we have

$$
\begin{aligned}
\mathbb{E}\left[(M_{t_i})^2\right] &= \frac{1}{\epsilon^4}\mathbb{E}\left[\left(\int_{t_{i-1}}^{t_i}\int_u^{t_i}\sum_{m=0}^{\infty}\frac{1}{m!}\left(\frac{t-u}{\epsilon^2}\right)^m \mathcal{L}_0^{(m)}(\nabla_y f\beta)(y_u)dtdV_u\right)^2\right] \\
&= \frac{1}{\epsilon^4}\mathbb{E}\left[\left(\int_{t_{i-1}}^{t_i}\sum_{m=0}^{\infty}\frac{1}{(\epsilon^2)^m m!}\left(\int_u^{t_i}(t-u)^m\, dt\right)\mathcal{L}_0^{(m)}(\nabla_y f\beta)(y_u)dV_u\right)^2\right] \\
&= \frac{1}{\epsilon^2}\mathbb{E}\left[\left(\int_{t_{i-1}}^{t_i}\sum_{m=0}^{\infty}\frac{(t_i-u)^{m+1}}{(\epsilon^2)^{m+1}(m+1)!}\mathcal{L}_0^{(m)}(\nabla_y f\beta)(y_u)dV_u\right)^2\right] \\
&= \frac{1}{\epsilon^3}\mathbb{E}\left[\int_{t_{i-1}}^{t_i}\left(\left(e^{\frac{(t_i-u)}{\epsilon^2}\mathcal{L}_0}-1\right)\left(\mathcal{L}_0^{(-1)}\nabla_y f\beta\right)(y_u)dV_u\right)^2\right] \\
&= \frac{1}{\epsilon^3}\mathbb{E}\left[\int_{t_{i-1}}^{t_i}\frac{(t_i-u)^2}{\epsilon^4}(\nabla_y f\beta)(y_u)^2du\right] \\
&= \frac{1}{\epsilon^7}\int_{t_{i-1}}^{t_i}(t_i-u)^2\mathbb{E}\left[(\nabla_y f\beta)(y_u)^2\right]du \\
&\leq \frac{C^2}{\epsilon^6}\int_{t_{i-1}}^{t_i}(t_i-u)^2du
\end{aligned}
$$

$$=\frac{C^2\delta^3}{3\epsilon^7}.$$

where in the fourth line we have applied Itô Isometry and Definition C.1. Also, for the inequality we have used the fact that $\left(\mathbb{E}\left[(\nabla_y f\beta)(y_u)^2\right]\right)^{1/2} \leq C$ which comes from the fact that the functions f and β by assumption are uniformly bounded. $\square$

For simplicity, in what follows sometimes we will use the following notation

$$\|\cdot\|_p := (\mathbb{E}|\cdot|^p)^{1/p}, \qquad p \in [1,\infty). \tag{C.7}$$

Lemma C.5. *Let* $\lambda(y) = \epsilon \sum_{m=2}^{\infty} \left(\frac{\delta}{\epsilon^2}\right)^m \frac{\mathcal{L}_0^{(m)}\Phi(y)}{m!}$, *then* $\|\lambda(y)\|_2 \leq C\frac{\delta^2}{\epsilon^3}$.

Proof.

$$\lambda(y) \;=\; \frac{\delta^2}{2\epsilon^3}\mathcal{L}_0^{(m)}\Phi(y) + \epsilon \sum_{m=3}^{\infty} \left(\frac{\delta}{\epsilon^2}\right)^m \frac{\mathcal{L}_0^{(m)}\Phi(y)}{m!}.$$

Thus, from triangular inequality

$$\|\lambda(y)\|_2 \;=\; C\frac{\delta^2}{2\epsilon^3} + R(\epsilon,\delta),$$

where the highest order of $R(\epsilon,\delta)$ with respect to δ/ϵ is of $\mathcal{O}\left(\frac{\delta^3}{\epsilon^5}\right)$. $\square$

Lemma C.6. *Show that*

$$\begin{aligned} IIE_{T_1}^{n,\epsilon} : \;&=\; e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f(y) - e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f(y)\Phi_K\left(-e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f(y)\right) \\ &\qquad + \sigma_K^2\phi_K\left(e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f(y)\right) \\ &\leq\; e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f(y) + \sigma_K^2\phi_K\left(e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f(y)\right), \end{aligned}$$

where σ_K^2 *is the variance of* K *and* Φ_K, ϕ_K *is the cdf and pdf of* K *respectively.*

Proof. The proof is an immediate result from the fact that $e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0} f(y) > 0$ for every y and Lemma B.4. $\square$

Lemma C.7. *Show that*

$$\left(\mathbb{E}\left(e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0}f(y_t)\right)^2\right)^{1/2} \leq e^{\frac{C\delta}{\epsilon^2}}. \tag{C.8}$$

Proof. By the DefinitionC.1,

$$\begin{aligned}
\left\|e^{\frac{\delta}{\epsilon^2}\mathcal{L}_0}f(y)\right\|_2 &= \left\|\sum_{m=0}^{\infty}\left(\frac{\delta}{\epsilon^2}\right)^m \frac{\mathcal{L}_0^{(m)}}{m!}f(y)\right\|_2 \\
&\leq \sum_{m=0}^{\infty}\left(\frac{\delta}{\epsilon^2}\right)^m \frac{\left\|\mathcal{L}_0^{(m)}f(y)\right\|_2}{m!} \\
&= e^{\frac{C\delta}{\epsilon^2}},
\end{aligned}$$

as required. □

Lemma C.8. *Show that*

$$\left(\mathbb{E}\left(\int_{(i-1)\delta}^{i\delta} e^{\frac{(i\delta-u)}{\epsilon^2}\mathcal{L}_0}(\nabla_y f\beta)(y_u)dV_u\right)^2\right)^{1/2} \leq \sqrt{\delta} + \mathcal{O}\left(\frac{\delta^2}{\epsilon^2}\right). \tag{C.9}$$

Proof. Define $\psi := \nabla_y f\beta$, then by Itô Isometry

$$\begin{aligned}
\mathbb{E}\left[\left(\int_{(i-1)\delta}^{i\delta} e^{\frac{(i\delta-u)}{\epsilon^2}\mathcal{L}_0}\psi(y_u)dV_u\right)^2\right] &= \mathbb{E}\left[\int_{(i-1)\delta}^{i\delta}\left(e^{\frac{(i\delta-u)}{\epsilon^2}\mathcal{L}_0}\psi(y_u)\right)^2 du\right] \\
&= \int_{(i-1)\delta}^{i\delta}\mathbb{E}\left[\left(e^{\frac{(i\delta-u)}{\epsilon^2}\mathcal{L}_0}\psi(y_u)\right)^2\right]du \\
&\leq \int_{(i-1)\delta}^{i\delta} e^{C\frac{(i\delta-u)}{\epsilon^2}}du \quad \text{(Lemma C.7)} \\
&= \epsilon^2\frac{e^{\frac{2C\delta}{\epsilon^2}}-1}{2C} \\
&= \delta + R(\epsilon,\delta),
\end{aligned}$$

where $R(\epsilon,\delta)$ is of $\mathcal{O}\left(\frac{\delta^2}{\epsilon^2}\right)$. . □

Lemma C.9. *Show that*

$$E_{T_1}^{n\epsilon} = \lim_{\epsilon\to 0}\lim_{n\to\infty}\mathbb{E}\left[f(y)\sum_{k=2}^{n}\frac{(n+1-k)}{(n\epsilon)^2}e^{\frac{(k-1)\delta}{\epsilon^2}\mathcal{L}_0}f(y)\mathbf{1}_{\mathcal{C}}\left(f(y)\right)\right]$$

$$= \frac{1}{2}\mathbb{E}\left[f(y)\Phi(y)\right].$$

Proof. First notice that since $e^{\frac{i\delta}{\epsilon^2}\mathcal{L}_0}f$ is bounded $\forall i \in \{1, \ldots, k-1\}$ we can use the bounded convergence theorem (see Lemma A.5) to take the limit inside the expectation so that

$$E^{n\epsilon}_{T_1} = \mathbb{E}\left[f(y)\lim_{\epsilon\to 0}\lim_{n\to\infty}\sum_{k=2}^{n}\frac{(n+1-k)}{(n\epsilon)^2}e^{\frac{(k-1)\delta}{\epsilon^2}\mathcal{L}_0}f(y)\mathbf{1}_{\mathcal{C}}\left(f(y)\right)\right]$$

But,

$$\Sigma := \lim_{n\to\infty}\sum_{k=2}^{n}\frac{(n+1-k)}{(n\epsilon)^2}e^{\frac{(k-1)\delta}{\epsilon^2}\mathcal{L}_0}f(y) \tag{C.10}$$

is by Definition C.1 (recall that $\delta = T/n$ where for simplicity we have assumed $T = 1$) equal to

$$\begin{aligned} \Sigma &= \lim_{n\to\infty}\sum_{k=2}^{n}\frac{(n+1-k)}{(n\epsilon)^2}\sum_{m=0}^{\infty}\left(\frac{k-1}{n\epsilon^2}\right)^m\frac{\mathcal{L}_0^{(m)}f(y)}{m!} \\ &= \sum_{m=0}^{\infty}\frac{\mathcal{L}_0^{(m)}f(y)}{m!}\lim_{n\to\infty}\sum_{k=2}^{n}\frac{(n+1-k)}{(n\epsilon)^2}\left(\frac{k-1}{n\epsilon^2}\right)^m, \end{aligned} \tag{C.11}$$

but,

$$\lim_{n\to\infty}\sum_{k=2}^{n}\frac{(n+1-k)}{(n\epsilon)^2}\left(\frac{k-1}{n\epsilon^2}\right)^m = \frac{1}{(m+1)(m+1)\left(\epsilon^2\right)^{m+1}}. \tag{C.12}$$

Hence, Eq.(C.11) becomes

$$\begin{aligned} \Sigma &= \epsilon^2\sum_{m=0}^{\infty}\frac{\mathcal{L}_0^{(m)}f(y)}{(m+2)!\left(\epsilon^2\right)^{m+2}} \\ &\overset{\nu=m+2}{=} \epsilon^2\sum_{\nu=2}^{\infty}\frac{\mathcal{L}_0^{(\nu-2)}f(y)}{\nu!\left(\epsilon^2\right)^{\nu}} \\ &= \epsilon^2\sum_{\nu=0}^{\infty}\frac{\mathcal{L}_0^{(\nu-2)}f(y)}{\nu!\left(\epsilon^2\right)^{\nu}} - \epsilon^2\mathcal{L}_0^{(-2)}f(y) - \mathcal{L}_0^{-1}f(y) \\ &= \epsilon^2\left(e^{\frac{1}{\epsilon^2}\mathcal{L}_0} - \mathcal{L}_0^{(-2)}\right)f(y) - \mathcal{L}_0^{-)}f(y) \\ &\to -\mathcal{L}_0^{-1}f(y) \overset{\text{Eq.(5.4)}}{=} \Phi(y). \end{aligned}$$

Putting everything together we get that

$$E_{T_1}^{n,\epsilon} = \mathbb{E}\left[f(y)\Phi(y)\mathbf{1}_{\mathcal{C}}\left(f(y)\right)\right]$$

Finally, the fact that f that satisfies the centering condition (see Eq.(5.5)) suggests that

$$E_{T_1}^{n,\epsilon} = \frac{1}{2}\mathbb{E}\left[f(y)\Phi(y)\right],$$

as required. $\square$